

Tilburg University

Multivariate empirical processes

Einmahl, J.H.J.

Publication date:
1987

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Einmahl, J. H. J. (1987). *Multivariate empirical processes*. (CWI Tract; No. 32). Centrum voor Wiskunde en Informatica.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

MULTIVARIATE EMPIRICAL PROCESSES

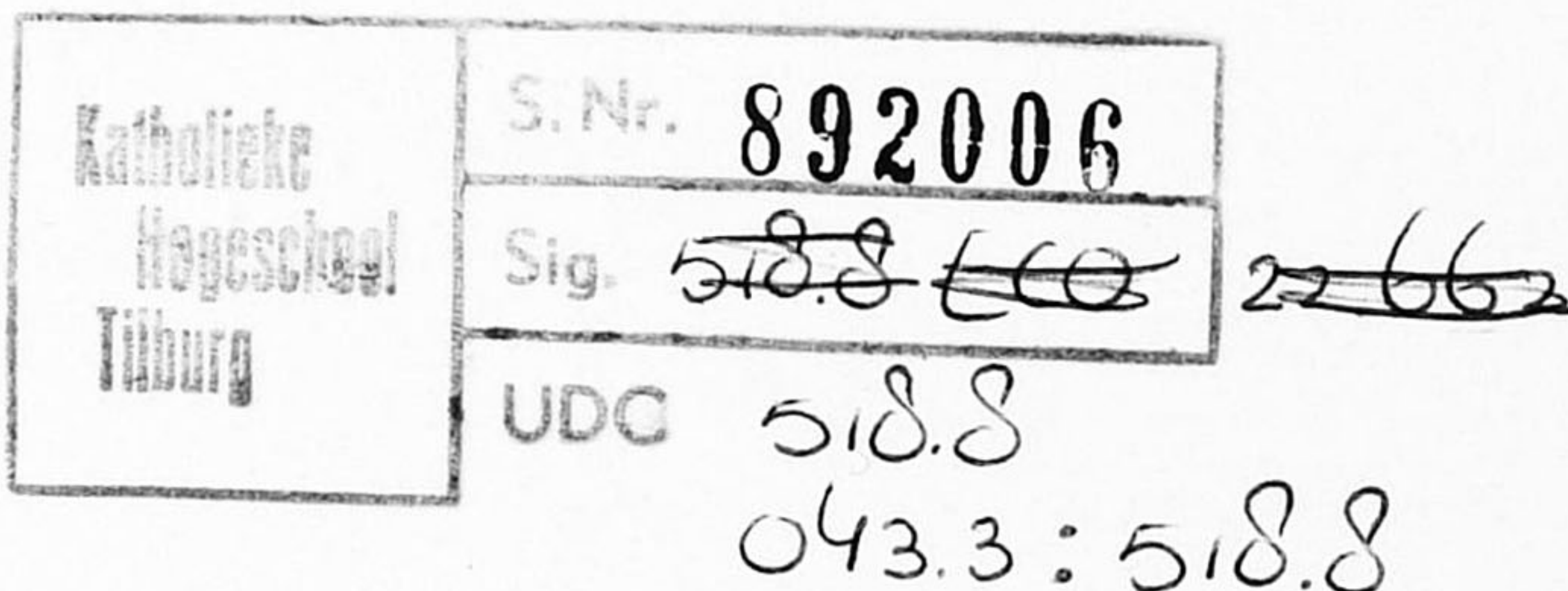
PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE
WISKUNDE EN NATUURWETENSCHAPPEN AAN DE
KATHOLIEKE UNIVERSITEIT TE NIJMEGEN
OP GEZAG VAN DE RECTOR MAGNIFICUS
PROF.DR. J.H.G.I GIESBERS
VOLGENS BESLUIT VAN HET COLLEGE VAN DEKANEN
IN HET OPENBAAR TE VERDEDIGEN
OP VRIJDAG 16 MEI 1986
DES NAMIDDAGS TE 2.00 UUR PRECIES

DOOR

JOHANNES HUBERTUS JACOB EINMAHL

GEBOREN TE HEERLEN



1986

PROMOTORES: PROF.DR. F.H. RUYMGAART

PROF.DR. D.M. MASON (U.S.A.)

800500	1983
1983	1983

Abstract

1. Introduction: This study aims to explore the relationship between the variables X and Y. The research is based on a sample of 100 individuals, with data collected over a period of 12 months. The primary objective is to determine the extent to which X influences Y, while controlling for other factors that may affect the outcome.

2. Methodology: The study employs a quantitative research design, utilizing a series of standardized questionnaires and statistical analysis. The data was collected through a structured survey, ensuring consistency and reliability. The statistical methods used include regression analysis to assess the strength and direction of the relationship between the variables.

3. Results: The findings indicate a positive correlation between X and Y. Specifically, as X increases, Y tends to increase as well. The regression analysis shows a significant positive slope, suggesting that the relationship is not merely coincidental but has a causal component. The results are supported by statistical significance tests, which confirm the validity of the findings.

4. Conclusion: Based on the data and analysis, it is concluded that there is a strong positive relationship between X and Y. This finding has important implications for understanding the underlying mechanisms at play. Further research is recommended to explore the long-term effects and to identify potential moderating factors that may influence the relationship.

5. Limitations: While the study provides valuable insights, it is not without limitations. The sample size, although adequate for the purposes of this study, may not be representative of the entire population. Additionally, the self-reported nature of the data may introduce some bias. Future studies should aim to address these limitations by using more diverse samples and objective measures.

Aan Fia

CONTENTS

CHAPTER 1. Introduction and basic facts	1
CHAPTER 2. Probability inequalities	7
1. Basic probability inequalities	7
2. Probability inequalities for weighted uniform empirical processes	12
3. Discussion and bibliography	22
CHAPTER 3. Criteria for weak convergence of the weighted multivariate uniform empirical process	24
1. Indexation by points	24
2. Indexation by rectangles	31
3. Discussion and bibliography	35
CHAPTER 4. Strong limit theorems for suprema of weighted multivariate uniform empirical processes	37
1. Suprema over the whole unit square: proper standardization is impossible	37
2. Suprema over the middle: a law of the iterated logarithm	48
3. Suprema over the tail: again laws of the iterated logarithm	53
4. Indexation by rectangles	62
5. Discussion and bibliography	64

CHAPTER 5. Limit theorems for the oscillation moduli of the multivariate uniform empirical process	67
1. The order of magnitude of the moments	67
2. The almost sure behaviour of the oscillation modulus ω_n	74
3. Discussion and bibliography	78
CHAPTER 6. Generalizations and applications	81
1. The negative part of the empirical process	81
2. Extension to a continuous density bounded away from 0 and ∞	84
3. Miscellaneous results	87
4. Some applications	90
REFERENCES	92
NOTATION INDEX	99
SUMMARY	100
SAMENVATTING	102
CURRICULUM VITAE	104

CHAPTER 1

INTRODUCTION AND BASIC FACTS

During the past decade very powerful methods have been discovered which have led to a wealth of new results in the theory of empirical processes. Many of these results concern the empirical process in a non-standard setting. One such modern and very fruitful approach is based on the concept of a Vapnik-Červonenkis indexing class. Without giving complete references, we mention Vapnik & Červonenkis (1971), Dudley (1978, 1984), Alexander (1982, 1984b), Devroye (1982), Gaenssler (1983) and Pollard (1984). Another powerful approach, applicable for independent identically distributed (iid) random variables (rv's), i.e. real valued random elements, is the method of strong and weak approximations. This approach is followed in Komlós, Major & Tusnády (1975a), Csörgö & Révész (1981), Csörgö, Csörgö, Horváth & Mason (1986) and Mason & van Zwet (1985). Simultaneously various authors developed the theory of empirical processes along the classical lines, i.e. they used in contrast to the approximation methodology a "direct approach", based on simple indexing sets. In this connection we cite Csáki (1977), Shorack & Wellner (1978, 1982, 1986), Wellner (1978), Stute (1982, 1984), Mason, Shorack & Wellner (1983) and Ruymgaart & Wellner (1984). For a review of the literature the reader is also referred to the survey paper by Gaenssler & Stute (1979).

This monograph deals exclusively with the classical approach to the study of weighted empirical processes based on iid random vectors. As indexing sets we either use quadrants (points) or rectangles with sides parallel to the coordinate axes. A number of optimal results is obtained, concerning weak convergence and strong limit theorems. The theory developed here is more or less self-contained and uses only a limited number of known results.

In order to be more explicit let us specify our setup and introduce

some notation. Throughout $\{X_i\}_{i=1}^{\infty}$ will denote a sequence of iid random vectors defined on a probability space (Ω, \mathcal{F}, P) taking values in the unit square I^d , with $I = [0,1]$, $d \in \mathbb{N}$. (We will adopt two dimensional terminology like "square" and "rectangle" although the dimension $d \in \mathbb{N}$ is arbitrary but fixed.) The common, but *arbitrary*, distribution function (df) of the random vectors is denoted by F . We assume for convenience that the X_i take their values in the open unit square $(0,1)^d$, so that in particular $\{t \in I^d : F(t) > 0\} \subset (0,1)^d$. If we need to display the coordinates of $t \in \mathbb{R}^d$ we write $t = \langle t_1, \dots, t_d \rangle = \langle t_j \rangle$ and if $t_j = \xi$ for all $1 \leq j \leq d$ we simply write $\langle \xi \rangle$. For $s, t \in \mathbb{R}^d$ we define $s \leq t$ to mean that $s_j \leq t_j$ for all $1 \leq j \leq d$; $s < t$ means that $s_j < t_j$ for all $1 \leq j \leq d$ and $\langle s_j \wedge t_j \rangle$ is denoted by $s \wedge t$.

As indexing sets we will either use quadrants of the form $[0, t_1] \times \dots \times [0, t_d]$, $t = \langle t_j \rangle \in I^d$, which are identified with the points t , or closed rectangles of the form $[s_1, t_1] \times \dots \times [s_d, t_d]$, $s = \langle s_j \rangle \leq t = \langle t_j \rangle$, $s, t \in I^d$, denoted by $R(s, t)$ or simply R . The class of all these closed rectangles is denoted by \bar{R} . Occasionally, we will also use the class R of all half-open rectangles $(s_1, t_1] \times \dots \times (s_d, t_d]$, $s < t$, $s, t \in I^d$, as an indexing class. Without confusion the rectangles in this class will also be denoted by $R(s, t)$ or R .

The *empirical distribution function* at stage $n \in \mathbb{N}$, the usual estimator of F for sample size n , is defined by

$$(1.1) \quad F_n(t) = \frac{1}{n} \# \{1 \leq i \leq n : X_i \leq t\}, \quad t \in I^d.$$

In order to keep our notation simple we use F and F_n to denote both the df and the corresponding probability measure; in the latter case we use $\{\cdot\}$ instead of (\cdot) , e.g. we write

$$(1.2) \quad F\{R\} = P(X_i \in R), \quad F_n\{R\} = \frac{1}{n} \# \{1 \leq i \leq n : X_i \in R\}, \quad R \in \bar{R} \text{ (or } R).$$

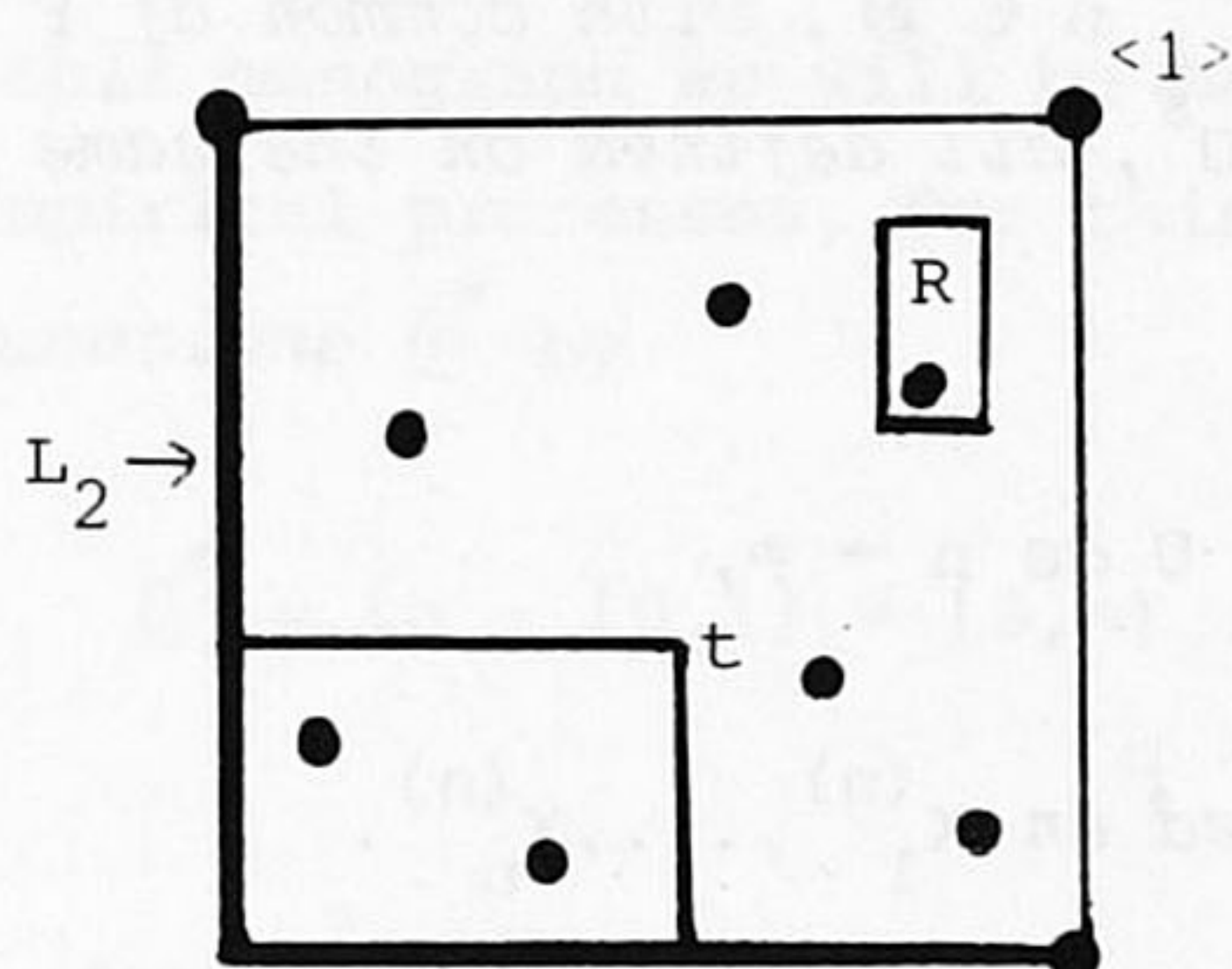
The *multivariate empirical process, indexed by points*, is now defined by

$$(1.3) \quad U_n^F(t) = n^{\frac{1}{2}} (F_n(t) - F(t)), \quad t \in I^d,$$

and the *multivariate empirical process, indexed by rectangles*, by

$$(1.4) \quad U_n^F\{R\} = n^{\frac{1}{2}} (F_n\{R\} - F\{R\}), \quad R \in \bar{R}.$$

We will also use the notation in (1.4) for $R \in \mathcal{R}$. Note that $U_n^F(\cdot)$ is tied down at the lower boundary $L_d = \{t \in I^d : t_1 \wedge \dots \wedge t_j = 0\}$ of I^d and at its upper vertex $\langle 1 \rangle$, meaning that $U_n^F(t) = 0$ if $t \in L_d$ or $t = \langle 1 \rangle$. This property follows from the fact that both F_n and F are 0 on L_d and 1 at $\langle 1 \rangle$.



$$\begin{aligned} d &= 2 & F_7(t) &= 2/7 \\ n &= 7 & F_7\{R\} &= 1/7 \end{aligned}$$

We now present four fundamental results for U_n^F . The first three results concern the central limit theorem, the last the law of the iterated logarithm. These results will be presented for an arbitrary df F on I^d , rather than for a continuous one as is usually the case in the literature. For $d \in \mathbb{N}$ let D_d be the generalization to dimension d of the well-known space $D = D[0,1]$ and let \mathcal{A}_d be the σ -algebra on D_d , generated by $J_1^{(d)}$, the generalization to dimension d of Skorohod's J_1 -topology on D . There exists a metric on D_d which generates $J_1^{(d)}$ and makes D_d a separable and complete metric space. For a description of this space and metric see Neuhaus (1971). Finally let ρ denote the supremum metric on D_d defined by

$$(1.5) \quad \rho(f_1, f_2) = \sup_{t \in I^d} |f_1(t) - f_2(t)|, \quad f_1, f_2 \in D_d.$$

FACT 1.1.A. (See e.g. Neuhaus (1971).)

$$(1.6) \quad U_n^F \xrightarrow{w} U^F \text{ in the } J_1^{(d)}\text{-topology on } D_d \text{ as } n \rightarrow \infty,$$

where \xrightarrow{w} denotes weak convergence and U^F is a Gaussian process with

$$(1.7) \quad EU^F(t) = 0 \text{ and } E(U^F(s)U^F(t)) = F(s \wedge t) - F(s)F(t), \quad s, t \in I^d.$$

Like the U_n^F , the limiting process U^F is tied down at the lower boundary $L_d = \{t \in I^d : t_1 \wedge \dots \wedge t_d = 0\}$ of I^d and at its upper vertex $\langle 1 \rangle$. Furthermore, if F is continuous, the sample paths of U^F are continuous as well.

FACT 1.1.B. (Skorohod (1956).) *There exists a triangular array of row independent random vectors $x_1^{(n)}, \dots, x_n^{(n)}$, $n \in \mathbb{N}$, with common df F and a Gaussian process \tilde{U}^F , equal in law to U^F , all defined on the same probability space, such that*

$$(1.8) \quad \sup_{t \in I^d} |\tilde{U}_n^F(t) - \tilde{U}^F(t)| \xrightarrow{\text{a.s.}} 0 \text{ as } n \rightarrow \infty,$$

where \tilde{U}_n^F is the empirical process based on $x_1^{(n)}, \dots, x_n^{(n)}$.

FACT 1.1.C. *For an A_d -measurable functional $h : D_d \rightarrow \mathbb{R}$ with $F(h \text{ is } \rho\text{-continuous at } U^F) = 1$ we have*

$$(1.9) \quad h(U_n^F) \xrightarrow{w} h(U^F) \text{ as } n \rightarrow \infty.$$

FACT 1.2. (Kiefer (1961).)

$$(1.10) \quad \limsup_{n \rightarrow \infty} \sup_{t \in I^d} |U_n^F(t)| / (\log \log n)^{\frac{1}{2}} \leq 2^{-\frac{1}{2}} \text{ a.s.}$$

If there exists a $t \in I^d$ with $F(t) = \frac{1}{2}$, then " \leq " may be replaced by " $=$ ".

Fact 1.1.A can be found in the literature for continuous df F , so that Fact 1.1.B follows for continuous F by the well-known Skorohod construction. To extend the latter result to arbitrary F we make use of the existence of a continuous df $C = C_F$ on I^d such that $F(t) = C(\langle F_j(t_j) \rangle)$, $t \in I^d$. This function is called the copula or dependence function; see e.g. Moore & Spruill (1975). It is easily seen that we may present $U_{n,F}$ in the following way:

$$(1.11) \quad U_n^F(t) = U_n^C(\langle F_j(t_j) \rangle), \quad t \in I^d.$$

(In the above the F_j are the marginals of F .) A little reflection now yields (1.8) for arbitrary F on I^d . Hence we also have (1.6) for arbitrary

F. It is also possible to give a more direct proof of Fact 1.1.A for arbitrary F with the aid of C_F , again, by generalizing the proof of Theorem 16.4 in Billingsley (1968) to dimension d . Finally note that Fact 1.1.C does not follow from the so called "continuous mapping theorem", since only continuity w.r.t. the supremum metric ρ is assumed. However, this fact follows immediately from Fact 1.1.B, because (1.8) is formulated in terms of ρ .

In this monograph we will be mainly concerned with weighted versions of the empirical processes. For this purpose let us define the class of *weight functions* Q^* by

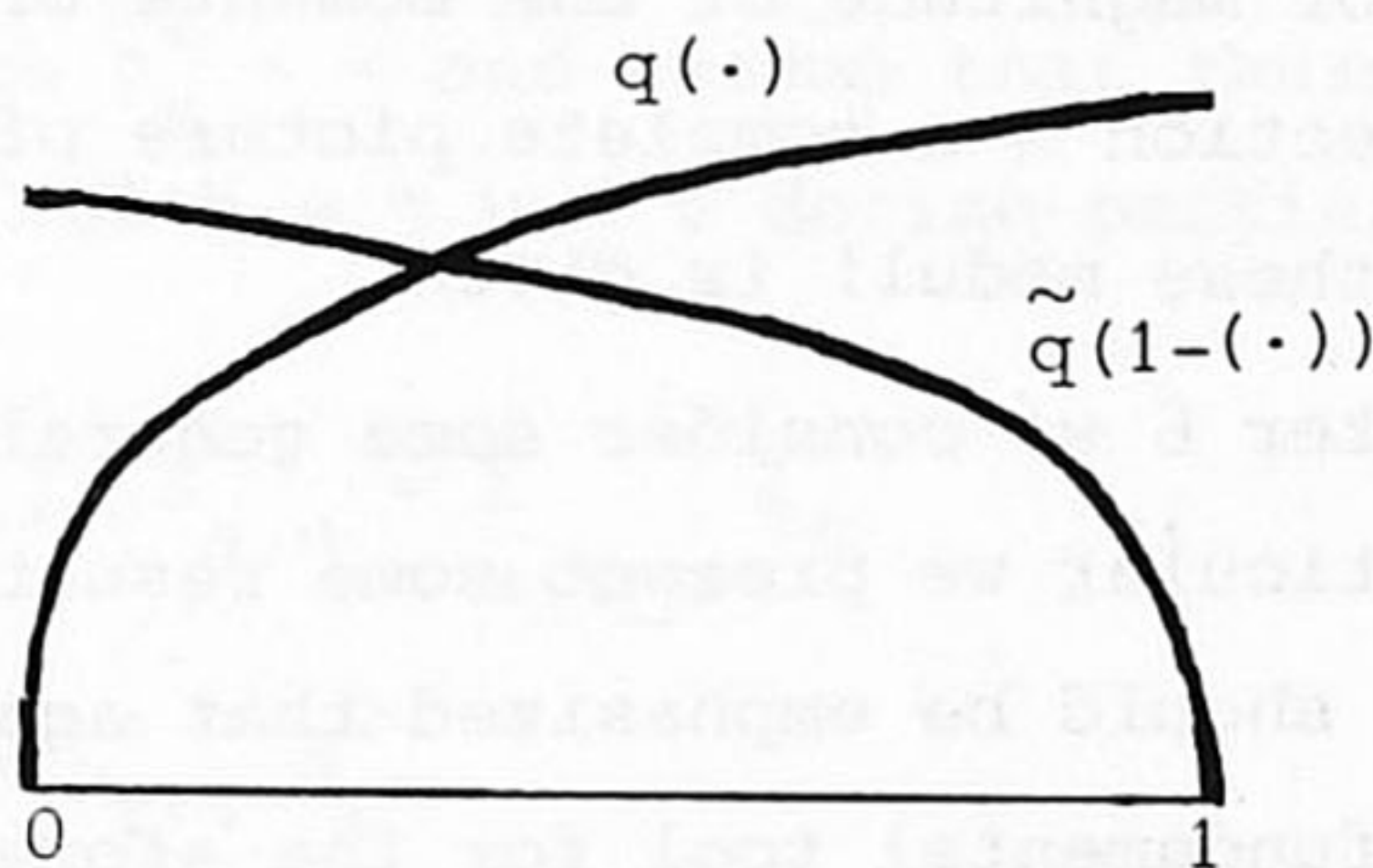
$$(1.12) \quad Q^* = \{q : [0,1] \rightarrow [0,\infty) : q \text{ continuous and non-decreasing, } q > 0 \text{ on } (0,1]\}.$$

Now for $q, \tilde{q} \in Q^*$ the *weighted multivariate empirical process, indexed by points*, is defined by

$$(1.13) \quad \frac{U_n^F(t)}{q(F(t))\tilde{q}(1-F(t))}, \quad t \in I^d,$$

and similarly the *weighted multivariate empirical process, indexed by rectangles*, by

$$(1.14) \quad \frac{U_n^F\{R\}}{q(F\{R\})\tilde{q}(1-F\{R\})}, \quad R \in \bar{R}.$$



(In this and other definitions we make the convention that $0/0 = 0$, $\sigma/0 = \infty$, for $\sigma > 0$, and $\sigma/0 = -\infty$, for $\sigma < 0$.) The most interesting weight functions turn out to be those $q \in Q^*$ with $q(0) = 0$, since U_n^F is tied down almost

surely, where F is 0 or 1. Notice that the weight function $q \equiv 1$, which yields the *unweighted* empirical process, is also an element of \mathcal{Q}^* .

We conclude this introductory chapter by giving a brief summary of the contents of this monograph. In Chapter 2, section 1, we will derive a probability inequality for the local behaviour of the empirical process for arbitrary df F . In Chapter 2, section 2, this inequality becomes a fundamental tool to derive global probability inequalities for weighted empirical processes under the assumption that the underlying distribution is uniform, i.e.

$$(1.15) \quad F(t) = |t|, \quad t \in I^d,$$

where $|t| = \prod_{j=1}^d t_j$ denotes the Lebesgue measure of $R(\langle 0 \rangle, t)$. *This uniformity assumption will be maintained throughout Chapters 3-5.* When this assumption holds, we will drop the superscript F from the notation, e.g. we write U_n instead of U_n^F ; U_n is called the *uniform empirical process* and U the *tied down Brownian sheet*.

In Chapter 3 we investigate the weak convergence of the weighted uniform empirical processes, indexed by points and rectangles respectively, for all possible weight functions. In Chapter 4 strong limit theorems for suprema of the absolute values of the weighted uniform empirical processes, mostly indexed by points, are obtained for weight functions of the form $q(\cdot) = (\cdot)^\alpha$, $\alpha \in [0, 1]$. These suprema are taken over various types of regions. Chapter 5 is concerned with results for two types of oscillation moduli of the unweighted uniform empirical process. More specifically in section 1 the order of magnitude of the moments of these oscillation moduli is obtained and in section 2 a complete picture of the almost sure behaviour of one of these moduli is given.

Finally in Chapter 6 we consider some generalizations and potential applications. In particular we present some results for $-U_n \vee 0$ and for U_n^F with smooth F . It should be emphasized that again the basic inequality in Chapter 2 is the fundamental tool for the aforementioned results. Among the fields of possible applications we mention extreme value theory, U -statistics, general spacings and density and regression function estimation.

Each of the Chapters 2-5 and sections 1 and 2 of Chapter 6 will be concluded by a discussion of the results, including the relevant references.

CHAPTER 2

PROBABILITY INEQUALITIES

In this chapter we collect the most important probability inequalities that we need for the proofs of our main results. In section 1 we first establish two versions of an inequality for partial sums of independent rv's indexed by elements of \mathbb{N}^d . With the aid of this inequality we obtain our basic inequality, concerning the fluctuation of the empirical process on a fixed arbitrary rectangle $R \in \mathcal{R}$. Typically in most applications this rectangle is small. In section 2 this basic inequality is used to establish sharp inequalities for the exceedance probability of the supremum of the weighted uniform empirical process along with two important maximal inequalities. These maximal inequalities are needed to obtain the almost sure results in Chapter 4.

1. Basic probability inequalities

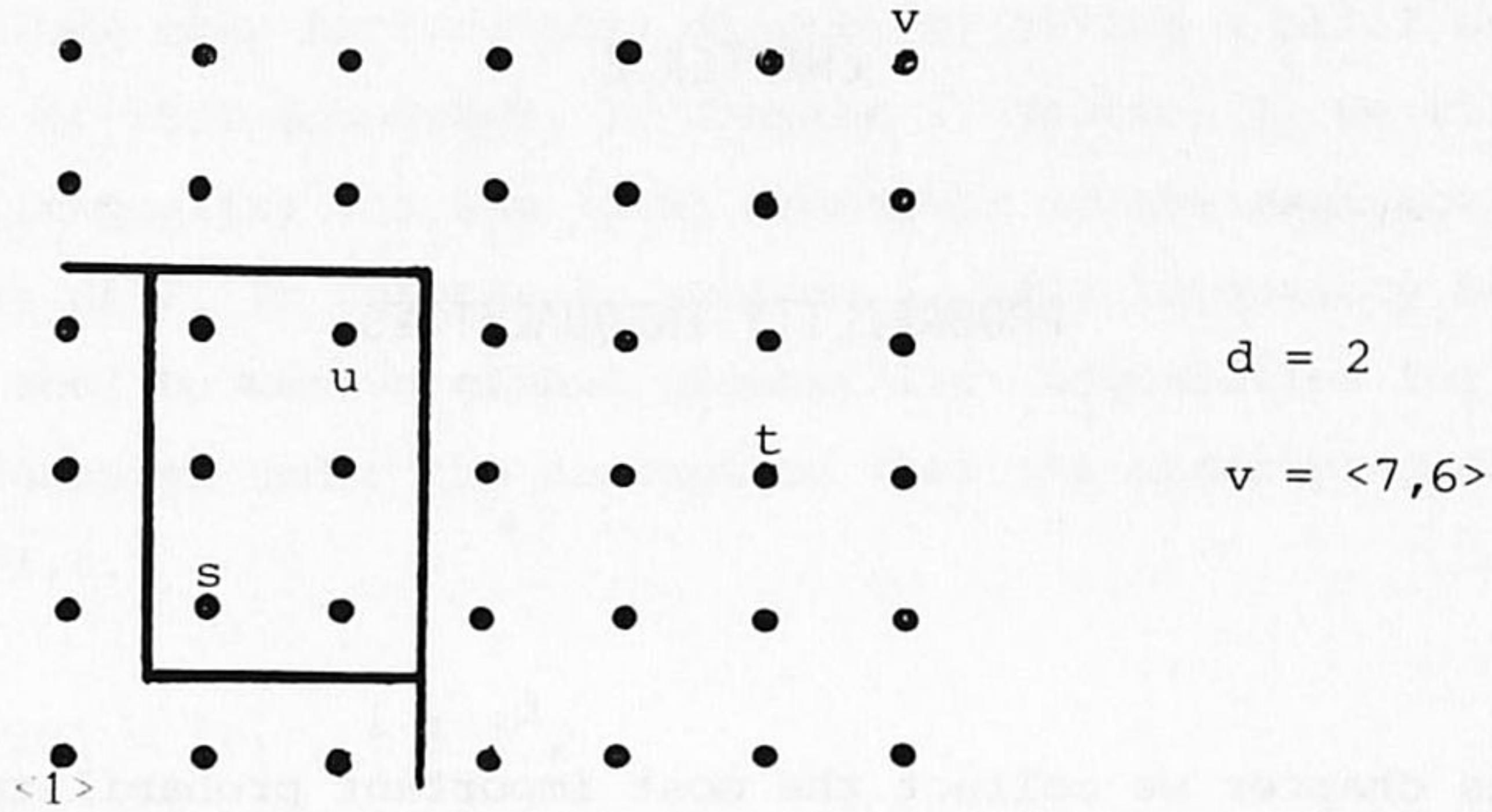
Let $d \in \mathbb{N}$ and fix $v \in \mathbb{N}^d$. For $t \in \mathbb{N}^d$ with $t \leq v$ let ξ_t be an rv with mean 0 and variance $\sigma_t^2 < \infty$ and assume that these rv's are independent. For arbitrary $s, u \in \mathbb{N}^d$ with $s \leq u \leq v$ define partial sums by

$$(2.1) \quad S_u = \sum_{t \leq u} \xi_t, \quad S_{s,u} = \sum_{s \leq t \leq u} \xi_t$$

and write $\sigma^2 = \text{Var } S_v = \sum_{t \leq v} \sigma_t^2$.

INEQUALITY 2.1. For all $\lambda \in \mathbb{R}$ we have

$$(2.2) \quad P(\max_{u \leq v} S_u \geq \lambda) \leq 2^d P(S_v \geq \lambda - d\sqrt{2\sigma^2}).$$



PROOF. Let $A = \{\max_{u \leq v} S_u \geq \lambda\}$ and $B = \{\max_{u \leq v; u_d = v_d} S_u \geq \lambda - \sqrt{2\sigma^2}\}$.
We first prove that

$$(2.3) \quad PA \leq 2PB.$$

Let $C = \{k \in \mathbb{N} : k \leq \prod_{j=1}^d v_j\}$, $D = \{t \in \mathbb{N}^d : t_j \leq v_j, \forall 1 \leq j \leq d\}$ and $g : C \rightarrow D$ a fixed bijective function with the property: $k < \ell$ implies $(g(k))_d \leq (g(\ell))_d$. Write $\tilde{g}(k) = \langle (g(k))_1, \dots, (g(k))_{d-1}, v_d \rangle$, $E_k = \{S_{g(k)} \geq \lambda\}$ and $A_k = E_k E_{k-1}^C \dots E_1^C$ ($k \in C$). Note that the A_k 's are pairwise disjoint and that $A = \bigcup_{k \in C} A_k$.

Now we have by the Chebyshev inequality and the independence of A_k and $\{S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}\}$ that

$$(2.4) \quad \begin{aligned} PA_k &\leq 2PA_k P(S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}) \\ &= 2P(A_k \cap \{S_{\tilde{g}(k)} - S_{g(k)} \geq -\sqrt{2\sigma^2}\}) \leq 2PA_k B. \end{aligned}$$

Hence

$$(2.5) \quad PA = \sum_{k \in C} PA_k \leq 2 \sum_{k \in C} PA_k B = 2PAB \leq 2PB,$$

which proves (2.3). It is now easily seen that the proof of Inequality 2.1 can be completed by an induction argument.

□

INEQUALITY 2.2. For all $\lambda \in \mathbb{R}$ we have

$$(2.6) \quad P(\max_{s \leq u \leq v} S_{s,u} \geq \lambda) \leq 2^{2d} P(S_v \geq \lambda - 2d\sqrt{2\sigma^2}).$$

PROOF. From the ξ_t , $t \in \mathbb{N}^d$, we derive a collection of rv's indexed by elements of \mathbb{N}^{2d} , which has the properties needed for application of Inequality 2.1. To distinguish the elements of \mathbb{N}^{2d} and \mathbb{N}^d we denote the elements of \mathbb{N}^{2d} by \bar{s} , \bar{t} etc. We consider $\bar{t} \leq \bar{w} = \langle v_1, \dots, v_d, v_1, \dots, v_d \rangle$. For \bar{t} with $\bar{t}_j + \bar{t}_{d+j} = v_j + 1$, $\forall 1 \leq j \leq d$, we define $\xi_{\bar{t}} = \xi_{\langle \bar{t}_1, \bar{t}_2, \dots, \bar{t}_d \rangle}$; $\xi_{\bar{t}} = 0$ otherwise. Now it is easily seen that

$S_{s,u} = S_{\langle u_1, \dots, u_d, v_1+1-s_1, \dots, v_d+1-s_d \rangle}$, which combined with an application of Inequality 2.1 (with d replaced by $2d$) yields

$$(2.7) \quad P(\max_{s \leq u \leq v} S_{s,u} \geq \lambda) \leq P(\max_{\bar{u} \leq \bar{w}} S_{\bar{u}} \geq \lambda) \\ \leq 2^{2d} P(S_{\bar{w}} \geq \lambda - 2d\sqrt{2\sigma^2}) = 2^{2d} P(S_v \geq \lambda - 2d\sqrt{2\sigma^2}).$$

□

Now we return to the empirical process as defined in (1.3) and (1.4). The next inequality is a useful relation between U_n^F and the process Z_n defined by

$$(2.8) \quad Z_n(t) = n^{-\frac{1}{2}}(N_n(t) - nF(t)), \quad t \in I^d,$$

where N_n is a Poisson process on I^d with $EN_n(t) = nF(t)$ for all $t \in I^d$. In the proof we shall use the fact that conditional on $N_n(\langle 1 \rangle) = n$ the processes U_n^F and Z_n are equal in law.

INEQUALITY 2.3. For any $R \in \mathcal{R}$ with $F\{R\} \leq \frac{1}{2}$, all $\lambda \in \mathbb{R}$ and either choice of sign we have

$$(2.9) \quad P(\sup_{S \in R} \pm U_n^F\{S\} \geq \lambda) \leq 2P(\sup_{S \in R} \pm Z_n\{S\} \geq \lambda),$$

where $S \in R$.

PROOF. Let $A_1 = \{\sup_{S \in R} \pm U_n^F\{S\} \geq \lambda\}$ and $A_2 = \{\sup_{S \in R} \pm Z_n\{S\} \geq \lambda\}$. Notice that $N_n\{R\}$ and $N_n\{R^c\}$ are independent. It follows that

$$\begin{aligned}
(2.10) \quad PA_1 &= P(A_2 \mid N_n(<1>) = n) \\
&= P(A_2 \cap \{N_n(<1>) = n\}) / P(N_n(<1>) = n) \\
&= \sum_{k=0}^n P(A_2 \cap \{N_n\{R\} = k\} \cap \{N_n\{R^C\} = n-k\}) / P(N_n(<1>) = n) \\
&= \sum_{k=0}^n (P(N_n\{R^C\} = n-k) / P(N_n(<1>) = n)) P(A_2 \cap \{N_n\{R\} = k\}) \\
&\leq \frac{P(N_n\{R^C\} = [nF\{R^C\}])}{P(N_n(<1>) = n)} \cdot PA_2,
\end{aligned}$$

because $P(Z = k) \leq P(Z = [\tau])$ if Z is a Poisson(τ) rv ($[\tau]$ = largest integer $\leq \tau$). An application of Stirling's formula completes the proof. \square

Note that the restriction $F\{R\} \leq \frac{1}{2}$ is used only at the end of the proof. It is easy to see that the number $\frac{1}{2}$ in the statement of Inequality 2.3 may be replaced by any number $a < 1$, provided only that we change the number 2 in (2.9) into a number depending on a . We will, however, only need this inequality for rectangles R with $F\{R\} \leq \frac{1}{2}$.

INEQUALITY 2.4. For any $R \in \mathcal{R}$, all $\lambda \in \mathbb{R}$ and either choice of sign we have

$$(2.11) \quad P(\sup_{S \subset R} \pm Z_n\{S\} \geq \lambda) \leq 2^{2d} P(\pm Z_n\{R\} \geq \lambda - 2d\sqrt{2F\{R\}}),$$

where $S \in \mathcal{R}$.

PROOF. For $k \in \mathbb{N}$ let $P_k \subset R$ be the partition of R consisting of 2^{kd} rectangles of equal Lebesgue measure with the property that the length of the i -th side of an element of P_k is 2^{-k} times the length of the i -th side of R . Observe now that the collection $\{\pm Z_n\{P\}\}_{P \in P_k}$ can be seen as a collection rv's $\{\xi_t\}_{t \leq v}$ (with the obvious indexation) to which Inequality 2.2 may be applied with $\sigma^2 = F\{R\}$.

We also have

$$(2.12) \quad \limsup_{k \rightarrow \infty}^* \pm Z_n\{S\} = \sup_{S \subset R} \pm Z_n\{S\} \quad \text{a.s.},$$

where \sup^* denotes the supremum over all $S \subset R$, $S \in \mathcal{R}$, such that S is a

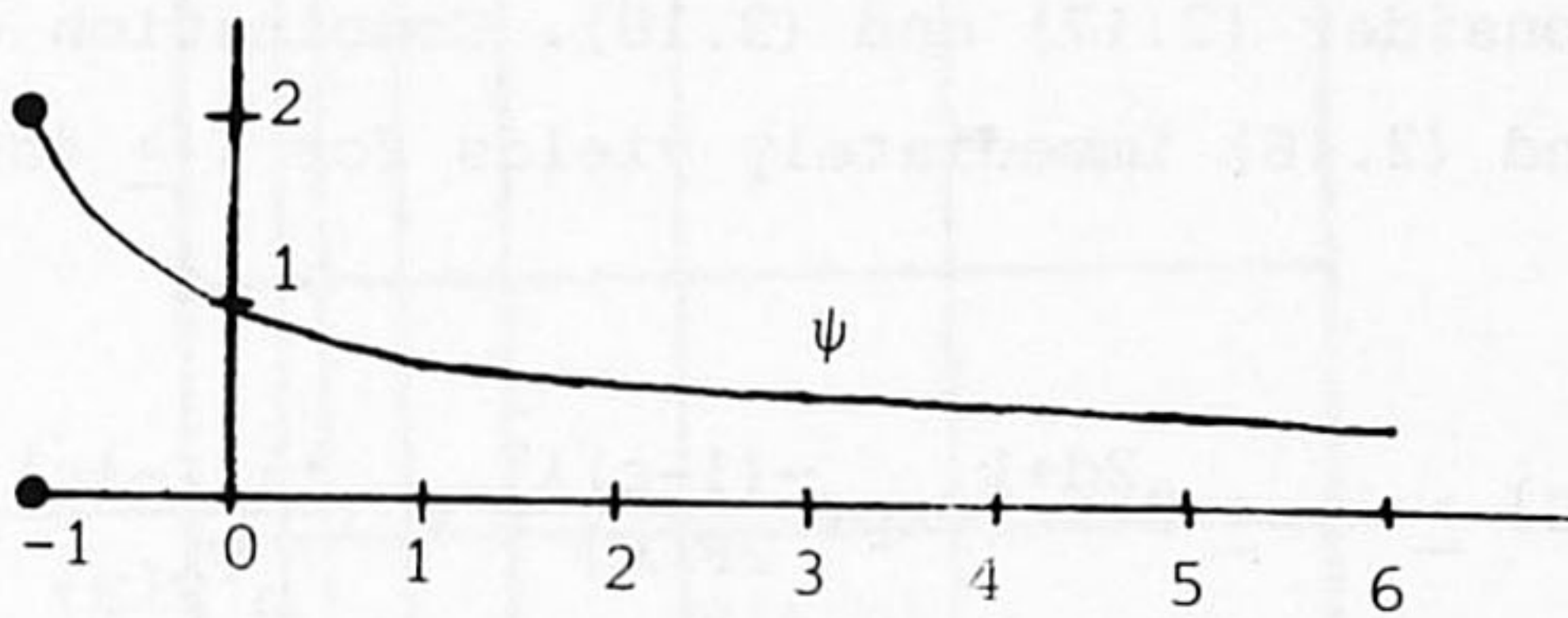
union of elements of P_k . Combination of the observation just above (2.12) and (2.12) itself yields for all $\lambda \in \mathbb{R}$

$$(2.13) \quad P(\sup_{S \subset R} \pm Z_n\{S\} > \lambda) \leq 2^{2d} P(\pm Z_n\{R\} \geq \lambda - 2d\sqrt{2F\{R\}}).$$

Since (2.13) holds for all $\lambda \in \mathbb{R}$ (2.11) easily follows. \square

The proof of the aforementioned basic inequality requires sharp bounds for the tail probabilities of a Poisson rv. For this purpose we define $\psi : [-1, \infty) \rightarrow [0, \infty)$ by

$$(2.14) \quad \begin{cases} \psi(\lambda) = 2\lambda^{-2} \int_0^\lambda \log(1+\sigma) d\sigma = 2\lambda^{-2} \{(1+\lambda)\log(1+\lambda) - \lambda\}, & \lambda \in (-1, 0) \cup (0, \infty); \\ \psi(0) = 1; \psi(-1) = 2. \end{cases}$$



This function has the following properties, see also Shorack & Wellner (1982, p. 641):

$$(2.15) \quad \begin{cases} \psi \text{ is decreasing and continuous,} \\ \psi(\lambda) \sim 2 \log \lambda / \lambda \text{ as } \lambda \rightarrow \infty, \\ \psi(\lambda) \geq 1/(1+\lambda/3) \quad \lambda \in [0, \infty), \\ (\cdot)\psi(\cdot) \text{ is increasing.} \end{cases}$$

Let $\tau > 0$, then we have for a Poisson(τ) rv Z and either choice of sign

$$(2.16) \quad P(\pm(Z - \tau) \geq \lambda) \leq \exp\left(\frac{-\lambda^2}{2\tau} \psi\left(\frac{\pm\lambda}{\tau}\right)\right), \quad \lambda \geq 0.$$

(For the minus part of (2.16) we also assume $\lambda \leq \tau$ because of the definition of ψ .) These bounds follow easily from computing the moment generating function and applying the Markov inequality.

INEQUALITY 2.5: basic inequality. Let $R \in \mathcal{R}$ with $0 < F\{R\} \leq \frac{1}{2}$. Then we have for any $\varepsilon \in (0,1)$

$$(2.17) \quad P(\sup_{S \in \mathcal{R}} U_n^F\{S\} \geq \lambda) \leq C^+ \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2F\{R\}} \psi\left(\frac{\lambda}{n^{\frac{1}{2}}F\{R\}}\right)\right), \quad \lambda \geq 0,$$

$$(2.18) \quad P(\sup_{S \in \mathcal{R}} -U_n^F\{S\} \geq \lambda) \leq C^- \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2F\{R\}} \psi\left(\frac{-(1-\varepsilon)\lambda}{n^{\frac{1}{2}}F\{R\}}\right)\right), \quad 0 \leq \lambda \leq n^{\frac{1}{2}}F\{R\},$$

$$(2.19) \quad P(\sup_{S \in \mathcal{R}} |U_n^F\{S\}| \geq \lambda) \leq C \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2F\{R\}} \psi\left(\frac{\lambda}{n^{\frac{1}{2}}F\{R\}}\right)\right), \quad \lambda \geq 0,$$

where $S \in \mathcal{R}$ and C^+ , C^- , $C \in (0, \infty)$ only depend on d and ε .

PROOF. Let us first consider (2.17) and (2.18). Combination of Inequality 2.3, Inequality 2.4 and (2.16) immediately yields for $\lambda \geq 4d\sqrt{2F\{R\}}/\varepsilon$ and either choice of sign

$$(2.20) \quad P(\sup_{S \in \mathcal{R}} \pm U_n^F\{S\} \geq \lambda) \leq 2^{2d+1} \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2F\{R\}} \psi\left(\frac{\pm\lambda(1-\frac{1}{2}\varepsilon)}{n^{\frac{1}{2}}F\{R\}}\right)\right).$$

(Again for the minus part we also assume $\lambda \leq n^{\frac{1}{2}}F\{R\}$.) For the other values of λ for which we need to prove (2.17) and (2.18) we can, of course, bound the probabilities by 1. From this remark and (2.20) we immediately obtain the bounds given in (2.17) and (2.18). Finally, observe that (2.19) follows from (2.17) and (2.18).

□

2. Probability inequalities for weighted uniform empirical processes

For the remainder of this chapter until the end of Chapter 5 we restrict ourselves to the *uniform* empirical process. Later on (e.g. in Chapter 3) it will become clear why some restriction w.r.t. the df has to be made. On the other hand, in Chapter 6 we will see in quite a few cases that the restriction to the uniform distribution is too severe, i.e. that we can prove the corresponding results in these cases for other df's as well.

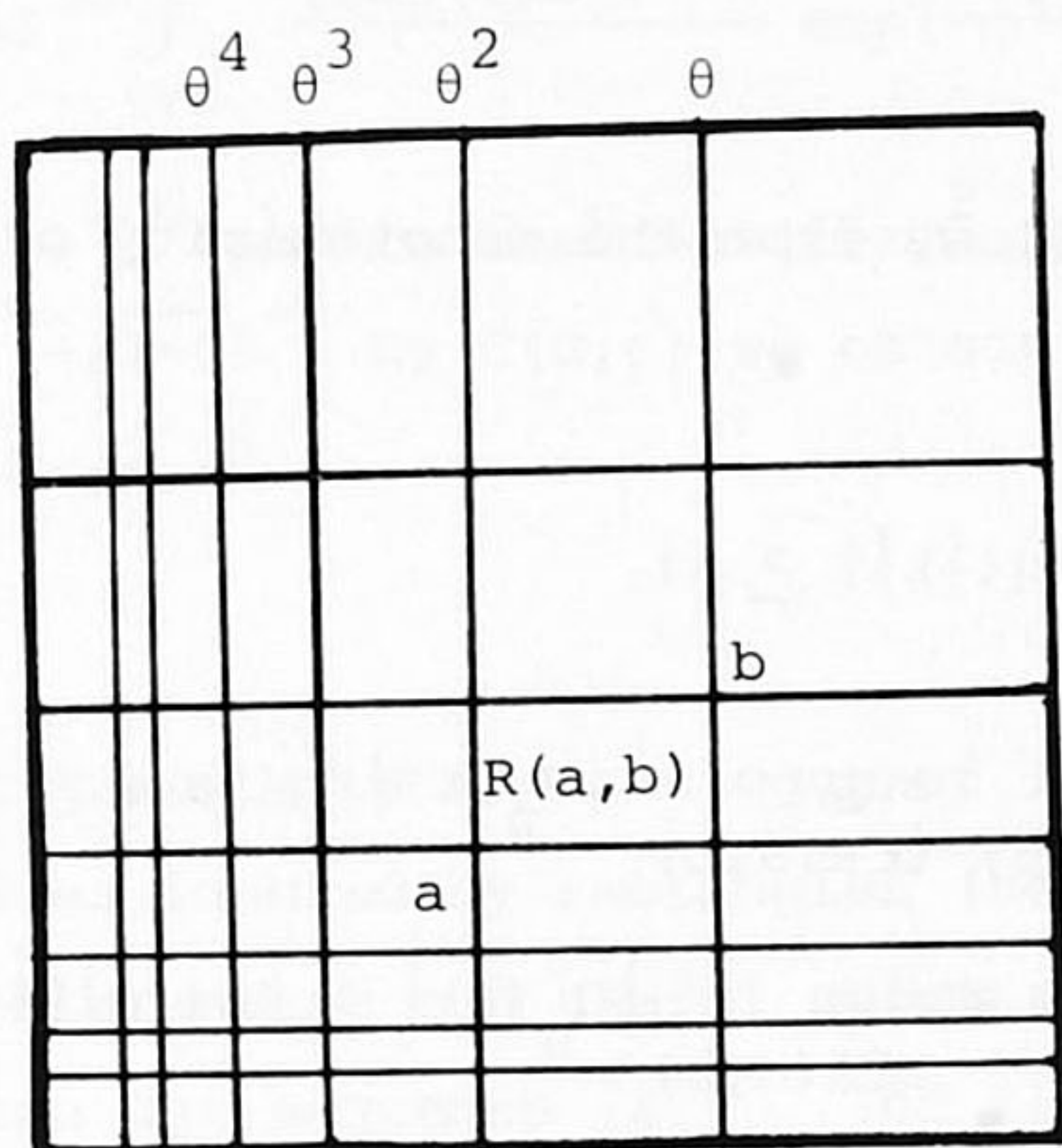
We introduce here a subclass of Q^* , see (1.12), defined by

$$(2.21) \quad Q = \{q \in Q^* : (\cdot)^{-\frac{1}{2}} q(\cdot) \text{ non-increasing on } (0,1]\}.$$

Next, we will use Inequality 2.5 to derive two essential inequalities for the weighted empirical process under the restriction that $q \in Q$ with $\tilde{q} = 1$. In Chapter 3 it will turn out that this is not a real restriction for our purposes, i.e. there we derive all the results for weight functions $q, \tilde{q} \in Q^*$.

For the proof of the first inequality we need a special countably infinite partition of $(0,1]^d$, that becomes arbitrarily fine near the lower boundary of the support of F which is equal to L_d , the lower boundary of I^d . For a fixed $\theta \in (0,1)$ this partition is the collection of rectangles

$$(2.22) \quad P(\theta) = \{R(\langle \theta^{k(j)} \rangle, \langle \theta^{k(j)-1} \rangle) \in R : \langle k(j) \rangle \in \mathbb{N}^d\}.$$



$$d = 2$$

$$P(\frac{2}{3})$$

For any $R(a,b) \in P(\theta)$ we have the useful property

$$(2.23) \quad \frac{|a|}{|b|} = \frac{\theta^{\sum_{j=1}^d k(j)}}{\theta^{\sum_{j=1}^d (k(j)-1)}} = \theta^d = \gamma \in (0,1);$$

note that γ is independent of the particular rectangle in the partition.

For any $0 < \alpha \leq \beta \leq 1$ let us introduce

$$(2.24) \quad P(\theta; \alpha, \beta) = \{R(a,b) \in P(\theta) : |b| \geq \alpha, |a| < \beta\},$$

consisting of all rectangles having a non-empty intersection with the set $\{t \in I^d : \alpha \leq |t| \leq \beta\}$. The inclusions

$$(2.25) \quad \{\alpha \leq |t| \leq \beta\} \subset \bigcup_{R \in \mathcal{P}(\theta; \alpha, \beta)} R \subset \{\gamma\alpha \leq |t| \leq \beta/\gamma\}$$

are immediate.

INEQUALITY 2.6. Let $q \in \mathcal{Q}$. Then we have for any $\varepsilon \in (0, 1)$ and $0 < \alpha \leq \beta \leq \frac{1}{2}(1-\varepsilon)$

$$(2.26) \quad \begin{aligned} & P(\sup_{\alpha \leq |t| \leq \beta} |U_n(t)|/q(|t|) \geq \lambda) \\ & \leq C \int_{(1-\varepsilon)\alpha}^{\beta/(1-\varepsilon)} \frac{(\log(1/\sigma))^{d-1}}{\sigma} \exp\left(-\frac{(1-\varepsilon)\lambda^2 q^2(\sigma)}{2\sigma}\right) \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right) d\sigma, \quad \lambda \geq 0, \end{aligned}$$

where $C = C(d, \varepsilon) \in (0, \infty)$.

PROOF. Let $\theta \in (0, 1)$. It follows from the monotonicity of q and from Inequality 2.5 that

$$(2.27) \quad \begin{aligned} & P(\sup_{\alpha \leq |t| \leq \beta} |U_n(t)|/q(|t|) \geq \lambda) \\ & \leq P(\max_{R(a,b) \in \mathcal{P}(\theta; \alpha, \beta)} \sup_{t \in R(a,b)} |U_n(t)|/q(|a|) \geq \lambda) \\ & \leq \sum_{R(a,b) \in \mathcal{P}(\theta; \alpha, \beta)} P(\sup_{t \in R(a,b)} |U_n(t)| \geq \lambda q(|a|)) \\ & \leq C \sum_{R(a,b) \in \mathcal{P}(\theta; \alpha, \beta)} \exp\left(-\frac{(1-\varepsilon)\lambda^2 q^2(|a|)}{2|b|}\right) \psi\left(\frac{\lambda q(|a|)}{n^{\frac{1}{2}}|b|}\right). \end{aligned}$$

In view of (2.23) and because $q \in \mathcal{Q}$ we may bound the first factor in the exponent in the right side of inequality (2.27) from below by

$$(2.28) \quad \frac{(1-\varepsilon)\lambda^2 q^2(|a|)}{2|b|} \geq \frac{\gamma(1-\varepsilon)\lambda^2 q^2(|t|)}{2|t|}, \quad \text{for } t \in R(a,b).$$

Using again $q \in \mathcal{Q}$ and the monotonicity of q and ψ , the second factor in the exponent in (2.27) may be bounded from below by

$$(2.29) \quad \psi\left(\frac{\lambda q(|a|)}{n^{\frac{1}{2}}|b|}\right) \geq \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right), \quad \text{for } R(a,b) \in \mathcal{P}(\theta; \alpha, \beta).$$

When we use, writing $|dt|$ for integration with respect to Lebesgue measure,

$$(2.30) \quad 1 = (1-\theta)^{-d} |b|^{-1} \int_{R(a,b)} |dt| \leq (1-\theta)^{-d} \int_{R(a,b)} 1/|t| |dt|$$

at the transition from summation to integration we find by combining (2.25), (2.27) - (2.29) that

$$(2.31) \quad \begin{aligned} & P(\sup_{\alpha \leq |t| \leq \beta} |U_n(t)| / q(|t|) \geq \lambda) \\ & \leq C(1-\theta)^{-d} \int_{\{\gamma\alpha \leq |t| \leq \beta/\gamma\}} \frac{1}{|t|} \exp\left(-\frac{\gamma(1-\varepsilon)\lambda^2 q^2(|t|)}{2|t|}\right) \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right) |dt|. \end{aligned}$$

The right side of this expression can be bounded from above, using the change of variables $\sigma = s_1 = |t|$, $s_2 = t_2, \dots, s_d = t_d$, by

$$(2.32) \quad C(1-\theta)^{-d} \int_{\gamma\alpha}^{\beta/\gamma} \frac{(\log(1/\sigma))^{d-1}}{\sigma} \exp\left(-\frac{\gamma(1-\varepsilon)\lambda^2 q^2(\sigma)}{2\sigma}\right) \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right) d\sigma.$$

Choosing $\theta = (1-\varepsilon)^{1/d}$ and relabeling $(1-\varepsilon)^2$ by $1-\varepsilon$ and then $C(d, 1-\sqrt{1-\varepsilon}) \cdot (1-\sqrt{1-\varepsilon})^{-d}$ by $C(d, \varepsilon)$ we obtain as an upper bound for (2.32) the expression in (2.26). □

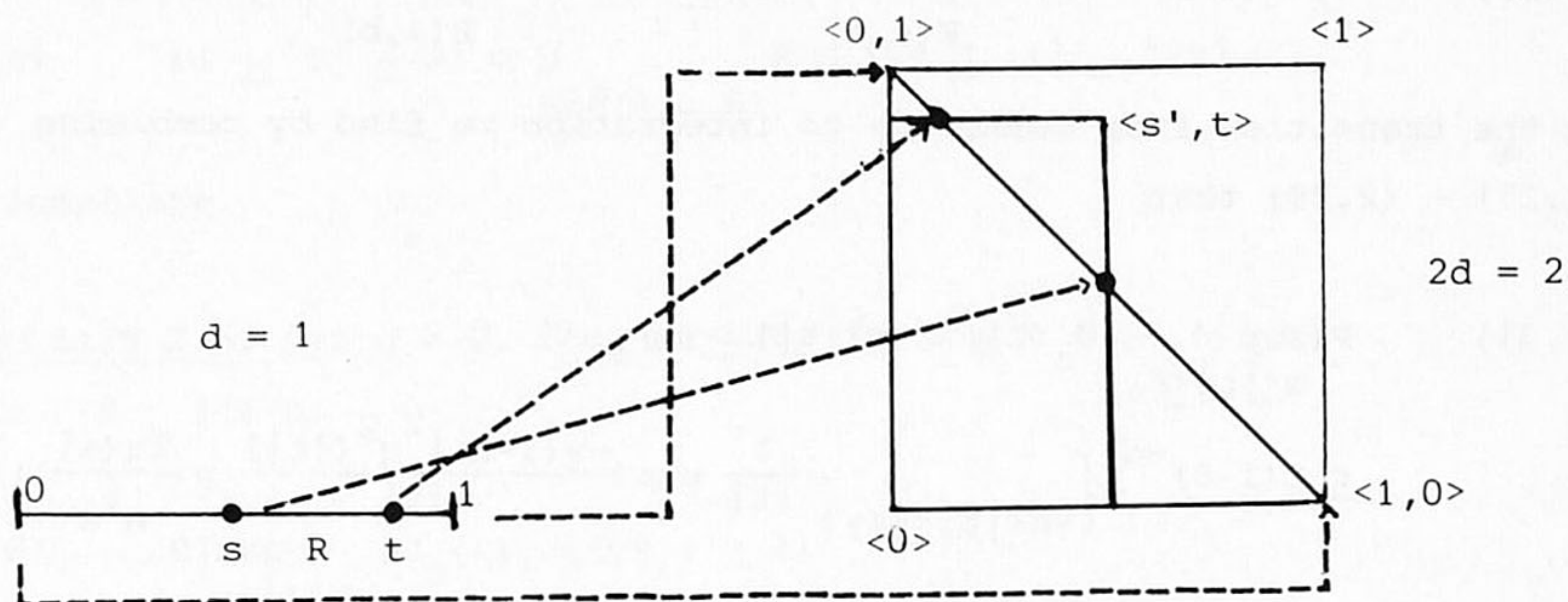
Our next inequality is the analogue of Inequality 2.6 for weighted empirical processes indexed by rectangles (see (1.14)). To accomplish this we will use an interesting and useful point representation for closed rectangles. Define the sequence $\{\bar{X}_i\}_{i=1}^\infty$ of iid rv's in I^{2d} by

$$(2.33) \quad \bar{X}_i = \langle 1 - X_{i1}, \dots, 1 - X_{id}, X_{i1}, \dots, X_{id} \rangle, \quad i \in \mathbb{N}.$$

The common df of the \bar{X}_i will be denoted by \bar{F} and the empirical df by \bar{F}_n . For $s \in I^d$ define $s' = \langle 1 \rangle - s = \langle 1 - s_j \rangle$. Observe that $\bar{F}(\langle s', t \rangle) = |t - s| = \prod_{j=1}^d (t_j - s_j)$ for $s, t \in I^d$ with $s \leq t$ and $\bar{F} = 0$ elsewhere. Let $\bar{U}_n(\langle s, t \rangle) = U_n^{\bar{F}}(\langle s, t \rangle)$, $s, t \in I^d$, be the empirical process on I^{2d} based on $\bar{X}_1, \dots, \bar{X}_n$. It is now easy to see that we have the following *point representation for the uniform empirical process indexed by rectangles*

$$(2.34) \quad \bar{U}_n(\langle s', t \rangle) = U_n\{R(s, t)\}, \quad s \leq t, \quad s, t \in I^d, \quad R \in \bar{R}.$$

Observe that a similar point representation is used in the proof of Inequality 2.2.



Now we need a covering of $I^{2d} \setminus \{\bar{F} = 0\}$, which will play a similar role as the partition in (2.22). This covering consists of a countably infinite number of rectangles $R \in \mathcal{R}$ and becomes arbitrarily fine near the lower boundary $\{<s,t> : s,t \in I^d, <s_j+t_j> \geq <1>, \exists_{1 \leq j \leq d} s_j+t_j = 1\}$ of the support of \bar{F} . (For instance, when $d = 1$ this boundary is the line segment joining $<0,1>$ and $<1,0>$.) The covering can be written as a product of a covering of the subset $A = \{<x,y> \in I^2 : x+y > 1\}$ of I^2 . Hence we have specified the covering completely if we define it on A .

For a fixed $\theta \in (\frac{1}{2}, 1)$ and $i \in \mathbb{N}$ let us first consider the set $A_i = \{<x,y> \in I^2 : \theta^i < x+y-1 \leq \theta^{i-1}\}$. It is easily seen that we can find a finite covering $\bar{P}_i'(\theta)$ of A_i with squares $R \in \mathcal{R}$ with the following properties:

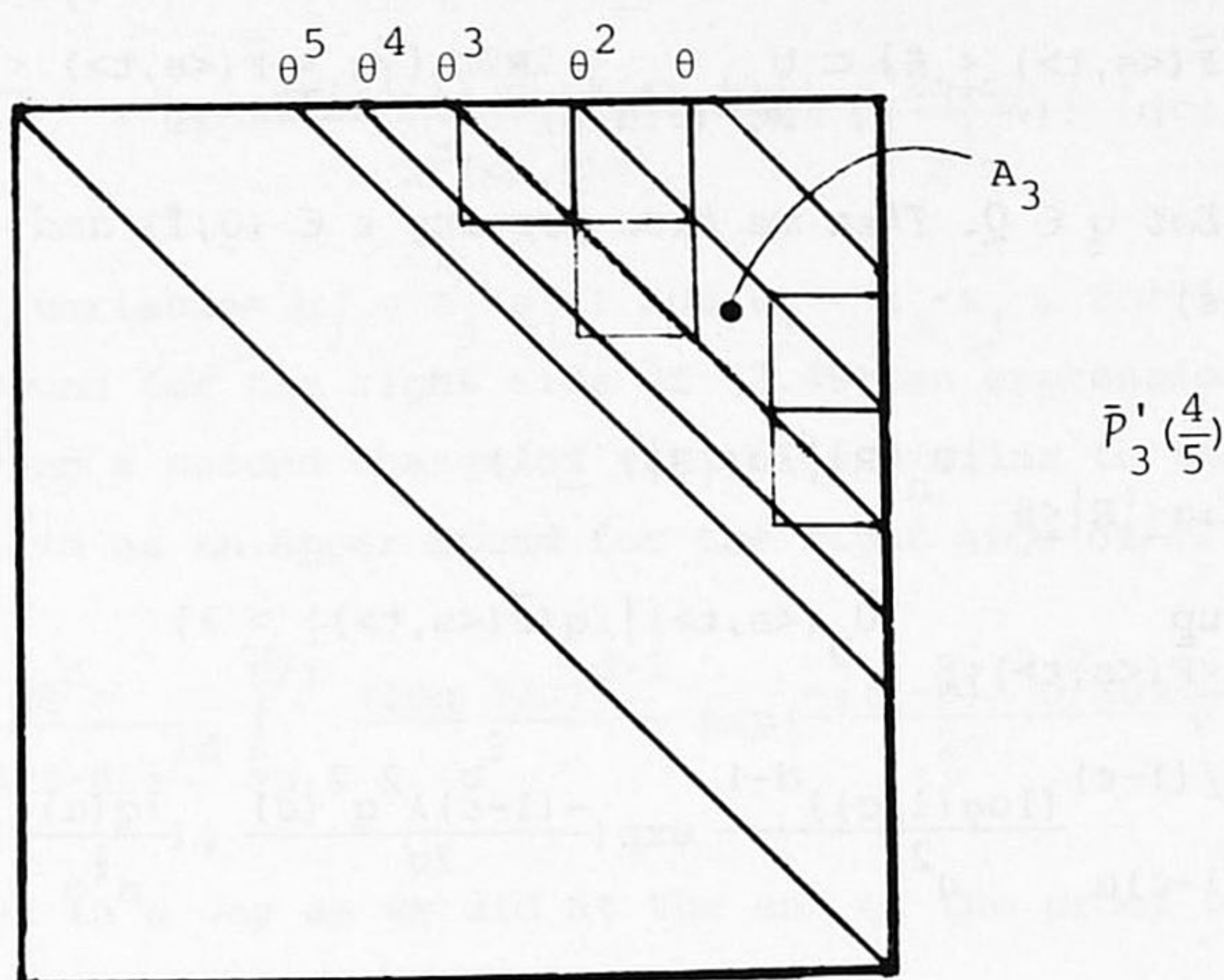
$$(2.35) \quad \bigcup_{R \in \bar{P}_i'(\theta)} R \subset \{<x,y> \in I^2 : \theta^{i+2} < x+y-1 \leq \theta^{i-2}\},$$

$$(2.36) \quad \text{each point of } I^2 \text{ is contained in at most 2 rectangles of } \bar{P}_i'(\theta),$$

$$(2.37) \quad \text{for each } R(<x_1,x_2>, <y_1,y_2>) \in \bar{P}_i'(\theta) \text{ we have} \\ (y_1-x_1)(y_2-x_2) = ((1-\theta)\theta^{n-1})^2 \geq (\theta(1-\theta)(y_1+y_2-1))^2.$$

The covering of A is now defined by

$$(2.38) \quad \bar{P}'(\theta) = \bigcup_{n=1}^{\infty} \bar{P}_i'(\theta);$$



from (2.35) and (2.36) it is clear that it has the property

(2.39) *each point of A is contained in at most 8 rectangles of $\bar{P}'(\theta)$.*

Finally we obtain the covering $\bar{P}(\theta)$ of $I^{2d} \setminus \{\bar{F} = 0\}$ by taking the Cartesian product of d times $\bar{P}'(\theta)$, where the coordinates s_j and t_j of $\langle s, t \rangle \in I^{2d}$ are taken pairwise together to form I^2 , $1 \leq j \leq d$. $\bar{P}(\theta)$ has the following properties:

(2.40) *for any $R(a, b) \in \bar{P}(\theta)$ holds $\frac{\bar{F}(a)}{\bar{F}(b)} \geq \theta^{4d} = \bar{\gamma} \in (0, 1)$,*

(2.41) *each point of $I^{2d} \setminus \{\bar{F} = 0\}$ is contained in at most 8^d rectangles of $\bar{P}(\theta)$,*

(2.42) $1 \leq (\theta(1-\theta))^{-2d} \int_R (\bar{F}(\langle s, t \rangle))^{-2} |d\langle s, t \rangle|$ for $R \in \bar{P}(\theta)$, $\langle s, t \rangle \in I^{2d}$.

Property (2.40) follows from (2.35), (2.41) is immediate from (2.39) and (2.42) from (2.37). Similar to $P(\theta; \alpha, \beta)$ in (2.24) we define for any $0 < \alpha \leq \beta \leq 1$

(2.43) $\bar{P}(\theta; \alpha, \beta) = \{R(a, b) \in \bar{P}(\theta) : \bar{F}(b) \geq \alpha, \bar{F}(a) < \beta\}$

and notice as in (2.25)

$$(2.44) \quad \{\alpha \leq \bar{F}(\langle s, t \rangle) \leq \beta\} \subset \bigcup_{R \in \bar{P}(\theta; \alpha, \beta)} R \subset \{\bar{\gamma}\alpha \leq \bar{F}(\langle s, t \rangle) \leq \beta/\bar{\gamma}\}.$$

INEQUALITY 2.7. Let $q \in \mathcal{Q}$. Then we have for any $\varepsilon \in (0, 1)$ and $0 < \alpha \leq \beta \leq \frac{1}{2}(1-\varepsilon)$

$$(2.45) \quad \begin{aligned} & P(\sup_{R \in \bar{P}(\theta; \alpha, \beta)} |U_n\{R\}|/q(|R|) \geq \lambda) \\ &= P(\sup_{\alpha \leq \bar{F}(\langle s, t \rangle) \leq \beta} |\bar{U}_n(\langle s, t \rangle)|/q(\bar{F}(\langle s, t \rangle)) \geq \lambda) \\ &\leq C \int_{(1-\varepsilon)\alpha}^{\beta/(1-\varepsilon)} \frac{(\log(1/\sigma))^{d-1}}{\sigma^2} \exp\left(-\frac{(1-\varepsilon)\lambda^2 q^2(\sigma)}{2\sigma}\right) \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right) d\sigma, \quad \lambda \geq 0, \end{aligned}$$

where $|R|$ denotes the Lebesgue measure of R and $C = C(d, \varepsilon) \in (0, \infty)$.

PROOF. The equality is immediate from (2.34), so we only have to prove the inequality. Let $\theta \in (\frac{1}{2}, 1)$. The same reasoning as in the proof of Inequality 2.6 yields

$$(2.46) \quad \begin{aligned} & P(\sup_{\alpha \leq \bar{F}(\langle s, t \rangle) \leq \beta} |\bar{U}_n(\langle s, t \rangle)|/q(\bar{F}(\langle s, t \rangle)) \geq \lambda) \\ &\leq C \sum_{R(a, b) \in \bar{P}(\theta; \alpha, \beta)} \exp\left(-\frac{(1-\varepsilon)\lambda^2 q^2(\bar{F}(a))}{2\bar{F}(b)}\right) \psi\left(\frac{\lambda q(\bar{F}(a))}{n^{\frac{1}{2}}\bar{F}(b)}\right); \end{aligned}$$

note that $a, b \in I^{2d}$. We also have

$$(2.47) \quad \frac{(1-\varepsilon)\lambda^2 q^2(\bar{F}(a))}{2\bar{F}(b)} \geq \frac{\bar{\gamma}(1-\varepsilon)\lambda^2 q^2(\bar{F}(\langle s, t \rangle))}{2\bar{F}(\langle s, t \rangle)}, \text{ for } \langle s, t \rangle \in R(a, b),$$

$$(2.48) \quad \psi\left(\frac{\lambda q(\bar{F}(a))}{n^{\frac{1}{2}}\bar{F}(b)}\right) \geq \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right), \text{ for } R(a, b) \in \bar{P}(\theta; \alpha, \beta).$$

Using (2.42) at the transition from summation to integration we find by combining (2.41), (2.44) and (2.46) - (2.48) that

$$(2.49) \quad P(\sup_{\alpha \leq \bar{F}(\langle s, t \rangle) \leq \beta} |\bar{U}_n(\langle s, t \rangle)|/q(\bar{F}(\langle s, t \rangle)) \geq \lambda)$$

$$\leq \frac{c8^d}{(\theta(1-\theta))^{2d}} \int_{\{\bar{\gamma}\alpha \leq \bar{F}(\langle s, t \rangle) \leq \beta/\bar{\gamma}\}} (\bar{F}(\langle s, t \rangle))^{-2} \\ \cdot \exp\left(\frac{-\bar{\gamma}(1-\varepsilon)\lambda^2 q^2(\bar{F}(\langle s, t \rangle))}{2\bar{F}(\langle s, t \rangle)} \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right)\right) |d\langle s, t \rangle|.$$

The change of variables $u_j = t_j + s_j - 1$ and $v_j = t_j - s_j$, for $1 \leq j \leq d$, yields as an upper bound for the right side of (2.49) an expression like (2.31).

Then by applying a second change of variables similar to the one below (2.31) we obtain as an upper bound for the right side of (2.49)

$$(2.50) \quad \frac{c8^d}{(\theta(1-\theta))^{2d}} \int_{\frac{\beta}{\bar{\gamma}}\alpha}^{\beta/\bar{\gamma}} \frac{(\log 1/\sigma)^{d-1}}{\sigma^2} \exp\left(\frac{-\bar{\gamma}(1-\varepsilon)\lambda^2 q^2(\sigma)}{2\sigma} \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right)\right) d\sigma.$$

Now we proceed in a way as we did at the end of the proof of Inequality 2.6 to obtain the expression in (2.45). □

For some purposes it is convenient to define for $v \in [0, \frac{1}{2}]$:

$$(2.51) \quad U_{n,v}(t) = n^v (F_n(t) - |t|), \quad t \in I^d.$$

Note that the processes $U_{n,\frac{1}{2}}$ and U_n are identical; depending on the context we will use one or the other of these notations.

INEQUALITY 2.8. Let $v \in [0, \frac{1}{2}]$. Then we have for any $\varepsilon \in (0, 1)$ and $0 < \alpha \leq \beta \leq \frac{1}{2}(1-\varepsilon)$

$$(2.52) \quad P\left(\sup_{\alpha \leq |t| \leq \beta} |U_{n,v}(t)| / |t|^{1-v} \geq \lambda\right) \\ \leq C \int_{(1-\varepsilon)\alpha}^{\beta/(1-\varepsilon)} \frac{(\log(1/\sigma))^{d-1}}{\sigma} \exp\left(-\frac{1}{2}(1-\varepsilon)\lambda^2 (n\sigma)^{1-2v} \psi(\lambda(n\alpha)^{-v})\right) d\sigma, \\ \lambda \geq 0,$$

where $C = C(d, \varepsilon) \in (0, \infty)$.

The proof is very much the same as the proof of Inequality 2.6 and will be omitted. The next inequality is the analogue for rectangles of Inequality 2.8. Its proof will not be given either, since it follows the lines of the proof of Inequality 2.7.

INEQUALITY 2.9. Let $v \in [0, \frac{1}{2}]$. Then we have for any $\varepsilon \in (0, 1)$ and $0 < \alpha \leq \beta \leq \frac{1}{2}(1-\varepsilon)$

$$\begin{aligned} & P(\sup_{R \in \mathbb{R}; \alpha \leq |R| \leq \beta} |U_{n,v}\{R\}| / |R|^{1-v} \geq \lambda) \\ & \leq C \int_{(1-\varepsilon)\alpha}^{\beta/(1-\varepsilon)} \frac{(\log(1/\sigma))^{d-1}}{\sigma^2} \exp(-\frac{1}{2}(1-\varepsilon)\lambda^2(n\sigma)^{1-2v} \psi(\lambda(n\alpha)^{-v})) d\sigma, \\ & \lambda \geq 0, \end{aligned}$$

where $C = C(d, \varepsilon) \in (0, \infty)$.

Finally we present two maximal inequalities which, when later combined with Inequality 2.8 or 2.9, will yield optimal almost sure results. For their proofs we need the following generalization of the "Lemma for events" in Loève (1977, p. 258):

LEMMA 2.1. Let I be a well-ordered (i.e. totally ordered and every subset has a smallest element) countable indexing set and for every $m \in I$ let A_m and B_m be events in some probability space. If for every $m \in I$, $A_m \cap (\bigcap_{k < m} A_k^C)$ and B_m are independent, then

$$\inf_{m \in I} P B_m \cdot P \bigcup_{m \in I} A_m \leq P \bigcup_{m \in I} (A_m \cap B_m).$$

PROOF. The proof is very much the same as the proof in Loève (1977). Hence it suffices to mention that $\bigcup_{m \in I} A_m = \bigcup_{m \in I} (A_m \cap (\bigcap_{k < m} A_k^C))$, which holds since I is a well-ordered set. \square

INEQUALITY 2.10. Let $v \in [0, \frac{1}{2}]$, $0 < \alpha \leq \beta \leq 1$, $\varepsilon \in (0, 1)$ and write $n_k = \lceil (1+\varepsilon/12)^k \rceil$, $k \in \mathbb{N}$. Then we have for all $k \in \mathbb{N}$ and for $\lambda > (2/\varepsilon)^v$

$$\begin{aligned} (2.53) \quad & P(\max_{n_k < n \leq n_{k+1}} \sup_{\alpha \leq |t| \leq \beta} |U_{n,v}(t)| / |t|^{1-v} \geq \lambda) \\ & \leq 2P(\sup_{\alpha \leq |t| \leq \beta} |U_{n_{k+1},v}(t)| / |t|^{1-v} \geq (1-\varepsilon)\lambda). \end{aligned}$$

PROOF. Let $\{r_i : i \in \mathbb{N}\}$ be an ordering of $\{t \in I^d : \alpha \leq |t| \leq \beta\} \cap \mathbb{Q}^d$ and write

$$(2.54) \quad S_n(t) = n(F_n(t) - |t|), \quad t \in I^d.$$

Since $U_{n,v} \in D_d$ we have that the first expression in (2.53), with " $\geq \lambda$ " replaced by " $> \lambda$ " (observe that this may be done without loss of generality, cf. the proof of Inequality 2.4), is less than or equal to

$$(2.55) \quad P(\sup_{n_k < n \leq n_{k+1}; i \in \mathbb{N}} |S_n(r_i)| / |r_i|^{1-v} > \lambda(n_k+1)^{1-v}).$$

For $n_k < n \leq n_{k+1}$ and $i \in \mathbb{N}$ set

$$(2.56) \quad A_{n,i} = \left\{ \frac{|S_n(r_i)|}{|r_i|^{1-v}} > \lambda(n_k+1)^{1-v} \right\}$$

and

$$(2.57) \quad B_{n,i} = \{ |S_{n_{k+1}}(r_i) - S_n(r_i)| \leq |r_i|^{1-v} \lambda(n_k+1)^{1-v} \cdot \frac{1}{2}\epsilon \}.$$

An application of Lemma 2.1 with the lexicographical ordering on $\{ \langle n, i \rangle : n_k < n \leq n_{k+1}, i \in \mathbb{N} \}$ yields

$$(2.58) \quad \inf_{n_k < n \leq n_{k+1}; i \in \mathbb{N}} PB_{n,i} \cdot \text{the probability in (2.55)} \\ \leq P \bigcup_{n_k < n \leq n_{k+1}; i \in \mathbb{N}} (A_{n,i} \cap B_{n,i}) \\ \leq P(\sup_{i \in \mathbb{N}} |S_{n_{k+1}}(r_i)| / |r_i|^{1-v} \geq (1 - \frac{1}{2}\epsilon) \lambda(n_k+1)^{1-v}).$$

It is easily seen that

$$(2.59) \quad P(\sup_{i \in \mathbb{N}} |S_{n_{k+1}}(r_i)| / |r_i|^{1-v} \geq (1 - \frac{1}{2}\epsilon) \lambda(n_k+1)^{1-v}) \\ \leq P(\sup_{\alpha \leq |t| \leq \beta} |U_{n_{k+1},v}(t)| / |t|^{1-v} \geq (1 - \epsilon) \lambda).$$

Hence we need only prove that $PB_{n,i} \geq \frac{1}{2}$ for all $n_k < n \leq n_{k+1}$ and all $i \in \mathbb{N}$. By elementary considerations it can be shown that for $1 \leq \alpha \leq 2$

$$(2.60) \quad E(|B(n,p) - np|^\alpha) \leq 3np,$$

where $B(n,p)$ is a binomial rv with parameters $n \in \mathbb{N}$ and $p \in (0,1)$.

Applying the Markov inequality with $\alpha = 1/(1-v)$ yields by the choice of λ

$$(2.61) \quad \text{PB}_{n,i}^c \leq \frac{3(n_{k+1} - (n_k + 1)) |r_i|}{(\lambda \varepsilon / 2)^{1/(1-\nu)} (n_k + 1) |r_i|} \leq \frac{1}{2} \varepsilon \left(\frac{2}{\lambda \varepsilon} \right)^{1/(1-\nu)} \leq \frac{1}{2}.$$

□

It should be noted that Inequality 2.10 and its proof remain valid, mutatis mutandis, for any underlying df F . Hence, choosing in particular $F = \bar{F}$, we have in view of (2.34)

INEQUALITY 2.11. Let $\nu \in [0, \frac{1}{2}]$, $0 < \alpha \leq \beta \leq 1$, $\varepsilon \in (0, 1)$ and write $n_k = \lceil (1 + \varepsilon/12)^k \rceil$, $k \in \mathbb{N}$. Then we have for all $k \in \mathbb{N}$ and for $\lambda > (2/\varepsilon)^\nu$

$$\begin{aligned} & P(\max_{n_k < n \leq n_{k+1}} \sup_{R \in \bar{R}; \alpha \leq |R| \leq \beta} |U_{n,\nu}\{R\}| / |R|^{1-\nu} \geq \lambda) \\ & \leq 2P(\sup_{R \in \bar{R}; \alpha \leq |R| \leq \beta} |U_{n_{k+1},\nu}\{R\}| / |R|^{1-\nu} \geq (1-\varepsilon)\lambda). \end{aligned}$$

3. Discussion and bibliography

Two papers by Ruymgaart & Wellner (1982, 1984) lay the foundation for this monograph and in particular for this chapter. In those papers Inequality 2.3, weaker versions of Inequalities 2.5 and 2.6 and the point representation for rectangles in (2.34) are established.

Inequality 2.1 with a different proof is contained in Klesov (1983); it generalizes an inequality in Kolmogorov (1929) w.r.t. the dimension. Inequality 2.2 appears to be new. Related inequalities can be found in Kiefer (1961, p. 651), Wichura (1969), Orey & Pruitt (1973, p. 142) and Paranjape & Park (1973). Apart from the usefulness of Inequalities 2.1 and 2.2 for this chapter, these inequalities might have a broader applicability. It is easy to see that Inequality 2.3 is true for functionals other than the supremum. Another version of this inequality can be found in Pyke & Shorack (1968); see also Donsker (1952). The proof of Inequality 2.4 is essentially contained in Orey & Pruitt (1973, p. 142).

Inequality 2.5, which is the final result of section 1, is an improvement of Theorem 1.4 in Stute (1984). It can also be compared to Theorem 3.1 in Alexander (1982) and Theorem 2.3 in Alexander (1984a) where the empirical process is assumed to be indexed by Vapnik-Červonenkis classes. It is a pleasant, though natural consequence of our approach, that the plus and minus parts can be kept separated. Hence also a sharp

bound for the minus part can be obtained; cf. the discussion above Inequality 1.2 in Shorack & Wellner (1982). Inequality 2.5 remains true for rectangles S and R that are both open or both closed. A proof follows by a slight modification of the proof of Inequality 2.4.

Most often, Inequality 2.5 will be applied to

$$P(\sup_{s \leq t} |U_n^F(s)| \geq \lambda),$$

which is a special case of the left side of (2.19), when R and S are chosen to be the quadrants $R(<0>, t)$ and $R(<0>, s)$. It is interesting to note that this version of Inequality 2.5 for quadrants could have been proved directly from Inequality 2.1, without the intervention of Inequality 2.2. An alternative proof of Inequality 2.5 for closed rectangles S and R can be obtained by combining Inequality 2.5 for quadrants with the point representation for rectangles in (2.34).

Inequalities 2.6 and 2.7 are generalizations w.r.t. the dimension of Inequalities 1.1 and 1.2 in Shorack & Wellner (1982). Inequality 2.6 can also be compared to Lemma 2.5 in Stute (1982) and to Theorem 2.1 in Alexander (1984b). The choice of the partition used in the proof on Inequality 2.6 is motivated by O'Reilly (1974), and the point representation in (2.34) by Kiefer & Wolfowitz (1958). Inequality 2.7 is a sharp version of Inequality 3.1 in Einmahl, Ruymgaart & Wellner (1984). Note that the exponent of σ in the integrand in (2.26) is equal to 1, whereas this exponent is equal to 2 in (2.45). This suggests that it might be difficult to derive a result like Inequality 2.6 with no restriction on the underlying df. Inequality 2.8, along with a proof, is contained in Einmahl (1984a). Inequality 2.10 is a generalization of Lemma 3(ii) in Shorack (1980); cf. also Lemma 2.3 in James (1975) and the proof of Lemma 7.2 in Alexander (1984b). Lemma 7.1 in the latter paper can be compared to our Lemma 2.1.

CHAPTER 3

CRITERIA FOR WEAK CONVERGENCE OF THE
WEIGHTED MULTIVARIATE UNIFORM EMPIRICAL PROCESS

In this chapter we derive necessary and sufficient conditions on the weight functions $q \in Q^*$ for weak convergence of the weighted multivariate uniform empirical process, indexed either by points (section 1) or rectangles (section 2). To avoid some inessential problems (i.a. measurability problems) we actually consider the convergence in probability to zero of the suprema of the weighted differences of \tilde{U}_n and \tilde{U} , where \tilde{U}_n and \tilde{U} are obtained from the Skorohod construction, presented in Fact 1.1.B. Without confusion, we shall henceforth drop the symbol \sim from the notation. All our results are stated for weight functions in Q^* , i.e. we will not need additional monotonicity conditions on the weight functions. For an explanation of the importance of this last remark we refer to section 3.

1. Indexation by points

It is the purpose of this section to give necessary and sufficient conditions on the weight functions q and \tilde{q} in order that

$$(3.1) \quad \sup_{t \in I^d} \frac{|U_n(t) - U(t)|}{q(|t|)\tilde{q}(1-|t|)} \xrightarrow{P} 0 \text{ as } n \rightarrow \infty.$$

The subclasses of Q^* that will appear in our theorems are

$$(3.2) \quad Q_0 = \{q \in Q^* : \int_0^1 \sigma^{-1} \exp(-\lambda q^2(\sigma)/\sigma) d\sigma < \infty \text{ for all } \lambda > 0\},$$

$$(3.3) \quad Q_k = \{q \in Q^* : \frac{q(\sigma)}{(\sigma(\log(1/\sigma))^k)^{\frac{1}{2}}} \rightarrow \infty \text{ as } \sigma \downarrow 0\}, k \in \mathbb{N}.$$

THEOREM 3.1. Let $d \in \mathbb{N}$ and $q \in \mathcal{Q}^*$. Then we have

$$(3.4) \quad \sup_{t \in I^d} |U_n(t) - U(t)| / q(|t|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if $q \in \mathcal{Q}_{d-1}$.

PROOF. The theorem is well-known for $d = 1$; see O'Reilly (1974). Hence we assume $d \geq 2$. The notation

$$(3.5) \quad g(\sigma) = \frac{q(\sigma)}{(\sigma(\log(1/\sigma))^{d-1})^{\frac{1}{2}}}, \quad \sigma > 0,$$

will be used in both parts of the proof.

(\Leftarrow) Suppose that $q \in \mathcal{Q}_{d-1}$. Following Shorack & Wellner (1982, p. 649) we may and will assume without loss of generality that

$$(3.6) \quad g(\cdot) \leq (\log(1/(\cdot)))^{(d-1)/2} \text{ and } g \downarrow \text{ on } (0,1] \text{ (hence } q \in \mathcal{Q}).$$

For any $0 < \delta \leq \frac{1}{2}$ we have

$$(3.7) \quad \sup_{t \in I^d} |U_n(t) - U(t)| / q(|t|) \leq \sum_{k=1}^5 Y_{nk},$$

where, with $\alpha_n = q^2(\frac{1}{n})$, $\beta_n = (d-1)!(n(\log n)^{d-1})^{-1}$ and $\gamma \in (0, \infty)$, the rv's Y_{nk} are given by

$$(3.8) \quad Y_{n1} = \sup_{0 \leq |t| \leq \beta_n / \gamma} |U_n(t)| / q(|t|),$$

$$(3.9) \quad Y_{n2} = \sup_{\beta_n / \gamma \leq |t| \leq \alpha_n} |U_n(t)| / q(|t|),$$

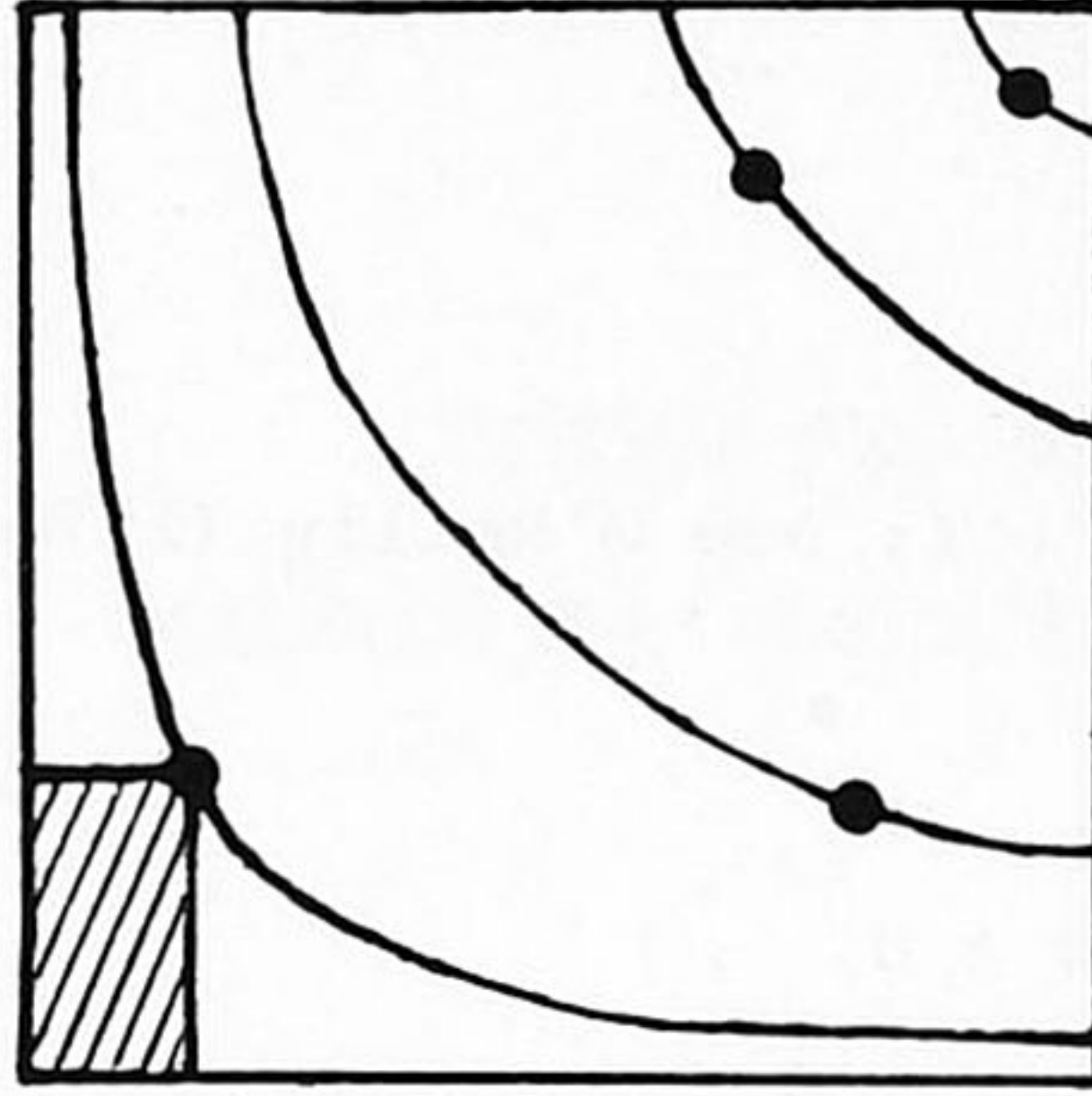
$$(3.10) \quad Y_{n3} = \sup_{\alpha_n \leq |t| \leq \delta} |U_n(t)| / q(|t|),$$

$$(3.11) \quad Y_{n4} = \sup_{0 \leq |t| \leq \delta} |U(t)| / q(|t|),$$

$$(3.12) \quad Y_{n5} = \sup_{t \in I^d} |U_n(t) - U(t)| / q(\delta).$$

It will be shown that for any $\varepsilon > 0$ and each $k = 1, \dots, 5$ there exist $\gamma = \gamma(\varepsilon)$, $\delta = \delta(\varepsilon)$ and $N = N(\varepsilon) \in \mathbb{N}$ such that

$$(3.13) \quad P(Y_{nk} \geq \varepsilon) \leq \varepsilon \text{ for } n \geq N.$$



$$d = 2$$

$$n = 4$$

To show (3.13) for $k = 1$ let $|X|_{1:n}^{(n)} = \min\{|X_1^{(n)}|, \dots, |X_n^{(n)}|\}$. Note that $P(|X|_{1:n}^{(n)} \leq \beta_n/\gamma) \rightarrow 1 - \exp(-1/\gamma)$, as $n \rightarrow \infty$, since $-\log|X_1^{(n)}|$ is a gamma rv with density $f_d(x) = ((d-1)!)^{-1} x^{d-1} e^{-x} 1_{(0,\infty)}(x)$. Hence we have $\lim_{n \rightarrow \infty} P(|X|_{1:n}^{(n)} \leq \beta_n/\gamma) < \varepsilon$ for γ sufficiently large. Under the condition $\sup_{0 \leq |t| \leq \beta_n/\gamma} F_n(t) = 0$, which is fulfilled with probability $\geq 1 - \varepsilon$ for n sufficiently large by the remark just made, it is easy to see that

$$\begin{aligned} (3.14) \quad Y_{n1} &\leq n^{\frac{1}{2}} \sup_{0 \leq |t| \leq \beta_n/\gamma} |t|/q(|t|) \\ &\leq n^{\frac{1}{2}} (\beta_n/\gamma)^{\frac{1}{2}} \{g(\beta_n/\gamma) (\log n)^{\frac{1}{2}(d-1)}\}^{-1} < \varepsilon. \end{aligned}$$

Hence it follows that for n sufficiently large

$$(3.15) \quad P(Y_{n1} \geq \varepsilon) \leq \varepsilon.$$

We have from (3.6) that for any $\gamma_1 \in (0, \infty)$

$$\begin{aligned} (3.16) \quad P(Y_{n2} \geq \varepsilon) &\leq P\left(\sup_{\beta_n/\gamma \leq |t| \leq \alpha_n} |U_n(t)|/|t|^{\frac{1}{2}} \geq \varepsilon g(\alpha_n) (\log(1/\alpha_n))^{\frac{1}{2}(d-1)}\right) \\ &\leq P\left(\sup_{\beta_n/\gamma \leq |t| \leq \alpha_n} |U_n(t)|/|t|^{\frac{1}{2}} \geq \gamma_1 (\log n)^{\frac{1}{2}(d-1)}\right), \end{aligned}$$

for $n \geq N_1 = N_1(\gamma_1)$. Hence, applying Inequality 2.6 with $q(\cdot) = (\cdot)^{\frac{1}{2}}$ and

$\varepsilon = \frac{1}{2}$ and using the second expression in (2.15), we see that there exist $c_1, \dots, c_4 \in (0, \infty)$ such that the last expression in (3.16) is bounded from above by

$$(3.17) \quad c_1 (\log n)^d \exp(-c_2 \gamma_1^2 (\log n)^{d-1} \psi(c_3 \gamma_1 \gamma^{\frac{1}{2}} (\log n)^{d-1})) \\ \leq c_1 (\log n)^d \exp(-c_4 \gamma_1 \gamma^{-\frac{1}{2}} \log \log n) \leq \varepsilon,$$

provided γ_1 and n are chosen sufficiently large.

Inequality 2.6 (with $\alpha = \alpha_n$, $\beta = \delta$ and $\varepsilon = \frac{1}{2}$) may be directly applied to (3.13) with $k = 3$. The integral in the resulting upper bound decreases to 0 as $n \rightarrow \infty$ and then $\delta \downarrow 0$, since $q \in \mathcal{Q}_{d-1}$ implies that

$$(3.18) \quad \int_0^1 \frac{1}{\sigma^2} \exp\left(\frac{-\lambda q^2(\sigma)}{\sigma}\right) d\sigma < \infty \text{ for all } \lambda > 0;$$

see Shorack & Wellner (1982, (1.9), (1.15) and (1.26)).

According to Orey & Pruitt (1973, Theorem 2.2) the function λq is point upper class for U , for all $\lambda > 0$. This yields

$\sup_{0 \leq |t| \leq \delta} |U(t)|/q(|t|) \xrightarrow{\text{a.s.}} 0$ as $\delta \downarrow 0$, which entails (3.13) for $k = 4$. The validity of (3.13) for $k = 5$ is immediate from (1.8).

(\Rightarrow) Let β_n be as before. We obviously have

$$(3.19) \quad \sup_{t \in I_d} |U_n(t) - U(t)|/q(|t|) \\ \geq \sup_{0 \leq |t| \leq \beta_n} |U_n(t) - U(t)|/q(|t|) = Y.$$

From the remark below (3.13) it follows that $P(|X|_{1:n}^{(n)} \leq \beta_n) \rightarrow 1 - e^{-1} > 3/5$. Hence with probability larger than $3/5$ we have $\sup_{0 \leq |t| \leq \beta_n} F_n(t) \geq n^{-1}$ for all large n , which in turn implies

$$(3.20) \quad Y \geq \{n^{\frac{1}{2}}(n^{-1} - \beta_n)\} - \sup_{0 \leq |t| \leq \beta_n} |U(t)|/q(\beta_n) = Z.$$

Now by applying Theorem 2.2 in Orey & Pruitt (1973) with, in their notation, $\phi(\xi) = (3N \log \log \xi)^{\frac{1}{2}}$ we obtain with probability arbitrarily close to 1 (for n large enough)

$$\sup_{0 \leq |t| \leq \beta_n} |U(t)| \leq \beta_n^{\frac{1}{2}} (3d \log \log (1/\beta_n))^{\frac{1}{2}} \leq 1/(3n^{\frac{1}{2}}).$$

Hence we have with probability larger than $\frac{1}{2}$ for all large n

$$(3.21) \quad Z \geq (2n^{\frac{1}{2}}q(\beta_n))^{-1} \geq \{3((d-1)!)^{\frac{1}{2}}g(\beta_n)\}^{-1}.$$

The assumption that $\sup_{t \in I^d} |U_n(t) - U(t)|/q(|t|) \xrightarrow{p} 0$ as $n \rightarrow \infty$, jointly with (3.19) - (3.21) implies that $g(\beta_n) \rightarrow \infty$ as $n \rightarrow \infty$. But since q is non-decreasing it is easily shown that $g(\sigma) \rightarrow \infty$ as $\sigma \downarrow 0$, i.e. that $q \in Q_{d-1}$. \square

THEOREM 3.2. Let $d \in \mathbb{N}$ and $\tilde{q} \in Q^*$. Then we have

$$(3.22) \quad \sup_{t \in I^d} |U_n(t) - U(t)|/\tilde{q}(1-|t|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if $\tilde{q} \in Q_0$.

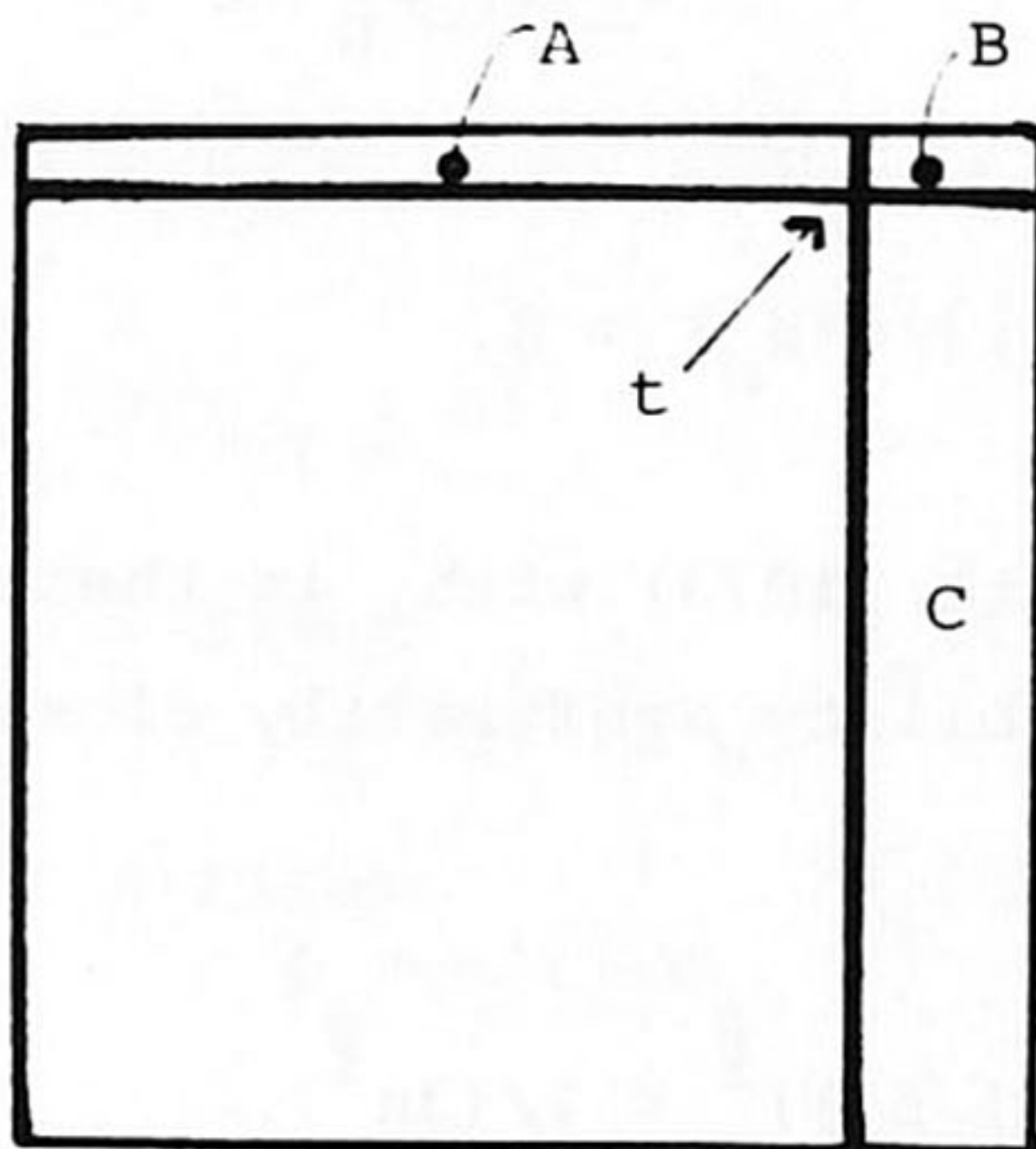
PROOF. Suppose $\tilde{q} \in Q_0$. Starting with the equalities

$$(3.23) \quad U_n(t) = -U_n\{R(<0>, t)^c\} \text{ and } U(t) = -U\{R(<0>, t)^c\}$$

we obtain using the union-intersection principle

$$(3.24) \quad |U_n(t) - U(t)| \leq \sum_{i \in I} |U_n\{R_i(t)\} - U\{R_i(t)\}|,$$

where the $R_i(t)$ are rectangles of the form $(s_1, 1] \times \dots \times (s_d, 1]$ with $s_j = t_j$ or $s_j = 0$ for all $1 \leq j \leq d$ and I is a finite indexing set. This yields



$$I = \{1, 2, 3\}$$

$$d = 2$$

$$R_1(t) = A \cup B$$

$$R_2(t) = B \cup C$$

$$R_3(t) = B$$

$$\begin{aligned}
 (3.25) \quad & \sup_{t \in I^d} |U_n(t) - U(t)| / \tilde{q}(1 - |t|) \\
 & \leq \sum_{i \in I} \sup_{t \in I^d} |U_n\{R_i(t)\} - U\{R_i(t)\}| / \tilde{q}(1 - |t|).
 \end{aligned}$$

It turns out to be convenient to split this sum into two parts. Define I_0 as the set of all $i \in I$ such that $R_i(t)$ equals $(0,1]^{j-1} \times (t_j,1] \times (0,1]^{d-j}$ for some $1 \leq j \leq d$. Write $I_1 = I \setminus I_0$. For $i \in I_0$ we have

$$\begin{aligned}
 (3.26) \quad & \sup_{t \in I^d} |U_n\{R_i(t)\} - U\{R_i(t)\}| / \tilde{q}(1 - |t|) \\
 & \leq \sup_{t \in I^d} |U_n\{R_i(t)\} - U\{R_i(t)\}| / \tilde{q}(|R_i(t)|).
 \end{aligned}$$

Application of Theorem 3.1 with $d = 1$ completes the proof for this part of the sum.

Now let $i \in I_1$. Define $D(R_i(t)) = \#\{j : R_i(t) \text{ depends on } t_j\}$. Suppose $D(R_i(t)) = \ell$, $2 \leq \ell \leq d$. By symmetry considerations studying

$$(3.27) \quad \sup_{t \in I^d} |U_n\{R_i(t)\} - U\{R_i(t)\}| / \tilde{q}(1 - |t|)$$

is equivalent to studying

$$(3.28) \quad \sup_{t \in I^d} |U_n(t') - U(t')| / \tilde{q}(1 - |\langle 1 \rangle - t|),$$

where t' is t restricted to I^ℓ in the way suggested above.

Defining $\xi = \max_{1 \leq j \leq d} t_j$ we have

$$(3.29) \quad \tilde{q}(1 - |\langle 1 \rangle - t|) \geq \tilde{q}(\xi),$$

and for small values of ξ

$$(3.30) \quad \tilde{q}(\xi) \geq \sqrt{\xi},$$

because $\tilde{q} \in \mathcal{Q}_0$, using an argument similar to Shorack & Wellner (1982, (a) on p. 648). Using $\xi^\ell \geq |t'|$ and $\ell \geq 2$ we have

$$(3.31) \quad \sqrt{\xi} \geq |t'|^\frac{1}{\ell}.$$

Theorem 3.1 implies that

$$(3.32) \quad \sup_{t' \in I} |U_n(t') - U(t')| / |t'|^{\frac{1}{d}} \xrightarrow{p} 0 \text{ as } n \rightarrow \infty.$$

Statement (3.32) when combined with (3.29) - (3.31) yields

$$(3.33) \quad \sup_{t \in I} |U_n(t) - U(t)| / \tilde{q}(1 - |t|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty.$$

This completes the if part of the proof.

The only if part is clear from the only if part in the onedimensional case by restricting the supremum e.g. to points of the form $t = \langle t_1, 1, \dots, 1 \rangle$.

□

COROLLARY 3.1. Let $d \in \mathbb{N}$ and $q, \tilde{q} \in Q^*$. Then we have

$$(3.34) \quad \sup_{t \in I^d} \frac{|U_n(t) - U(t)|}{q(|t|)\tilde{q}(1-|t|)} \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if both $q \in Q_{d-1}$ and $\tilde{q} \in Q_0$.

In Chapter 1, Fact 1.1.C, we saw that $h(U_n) \xrightarrow{w} h(U)$ for a functional $h : D_d \rightarrow \mathbb{R}$ which is A_d -measurable and for which $P(h \text{ is } \rho\text{-continuous in } U) = 1$. From Corollary 3.1 we can deduce a much stronger result. Before presenting it we need some notation. Let $q, \tilde{q} \in Q^*$. Define

$$(3.35) \quad q\tilde{q}D_d = \{g \in D_d : g(t) = q(|t|)\tilde{q}(1-|t|)f(t) \text{ for all } t \in I^d \text{ and } f \in D_d\}$$

and the metric $\rho_{q\tilde{q}}$ on $q\tilde{q}D_d$ by

$$(3.36) \quad \rho_{q\tilde{q}}(g_1, g_2) = \rho(g_1/q\tilde{q}, g_2/q\tilde{q}), \quad g_1, g_2 \in q\tilde{q}D_d,$$

where for $g \in q\tilde{q}D_d$, $g/q\tilde{q}$ is defined by

$$(3.37) \quad (g/q\tilde{q})(t) = g(t)/(q(|t|)\tilde{q}(1-|t|)) \text{ for } t \in I^d.$$

COROLLARY 3.2. Let $d \in \mathbb{N}$, $q \in \mathcal{Q}_{d-1}$, $\tilde{q} \in \mathcal{Q}_0$ and $h : D_d \rightarrow \mathbb{R}$ a functional which is A_d -measurable and for which $P(h \text{ is } \rho_{q\tilde{q}}\text{-continuous in } U) = 1$, then $h(U_n) \xrightarrow{w} h(U)$ as $n \rightarrow \infty$.

PROOF. We will only give a sketch of the proof because it is more or less standard. For $q \in \mathcal{Q}_{d-1}$ and $\tilde{q} \in \mathcal{Q}_0$ we have by Corollary 3.1 that $\rho_{q\tilde{q}}(U_n, U) \xrightarrow{p} 0$. Now h is A_d -measurable and $P(h \text{ is } \rho_{q\tilde{q}}\text{-continuous in } U) = 1$ imply $|h(U_n) - h(U)| \xrightarrow{p} 0$ as $n \rightarrow \infty$. This immediately yields $h(U_n) \xrightarrow{w} h(U)$ as $n \rightarrow \infty$. \square

2. Indexation by rectangles

At first sight one would expect that the purpose of this section, similarly to section 1, is to give necessary and sufficient conditions on the weight functions q and \tilde{q} in order that

$$\sup_{R \in \bar{R}} \frac{|U_n\{R\} - U\{R\}|}{q(|R|)\tilde{q}(1-|R|)} \xrightarrow{p} 0 \text{ as } n \rightarrow \infty.$$

It is easily seen, however, by extending an example in Shorack & Wellner (1982) to the multivariate case, that

$$(3.38) \quad \sup_{R \in \bar{R}; 0 < |R| \leq \varepsilon} |U_n\{R\}| / q(|R|) = \infty \text{ a.s., for all } 0 < \varepsilon < 1,$$

for any $q \in \mathcal{Q}^*$ with $q(0) = 0$. For this reason $|R|$ has to be bounded away from 0 when the weighted empirical process indexed by rectangles is studied.

LEMMA 3.1. In the setup of this chapter we have

$$(3.39) \quad \sup_{R \in \bar{R}} |U_n\{R\} - U\{R\}| \xrightarrow{\text{a.s.}} 0 \text{ as } n \rightarrow \infty.$$

PROOF. Since a half-open rectangle R is determined by 2^d points it is immediate from Fact 1.1.B that (3.39) holds with \bar{R} instead of \bar{R} . Now notice that for every rectangle $R \in \bar{R}$ a sequence of rectangles $\{R_i\}_{i=1}^{\infty}$, with $R_i \in \bar{R}$ can be found such that $U_n\{R_i\} - U\{R_i\} \xrightarrow{\text{a.s.}} U_n\{R\} - U\{R\}$ as $i \rightarrow \infty$. Hence (3.39) easily follows. \square

THEOREM 3.3. Let $d \in \mathbb{N}$, $q \in Q^*$ and $\gamma \in (0, \infty)$ fixed. Then we have

$$(3.40) \quad \sup_{R \in \bar{R}; \gamma \log n/n \leq |R| \leq 1} |U_n\{R\} - U\{R\}|/q(|R|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if $q \in Q_1$.

PROOF. (\Leftarrow) Suppose that $q \in Q_1$. As in the proof of Theorem 3.1 the notation

$$(3.41) \quad g(\sigma) = \frac{q(\sigma)}{(\sigma \log(1/\sigma))^{\frac{1}{2}}}, \quad \sigma > 0,$$

will be used and as in (3.6) we may and will assume without loss of generality that

$$(3.42) \quad g(\cdot) \leq (\log(1/(\cdot)))^{\frac{1}{2}} \text{ and } g \uparrow \text{ on } (0, 1] \text{ (hence } q \in Q).$$

For any $0 < \delta \leq \frac{1}{4}$ we have

$$(3.43) \quad \sup_{\gamma \log n/n \leq |R| \leq 1} |U_n\{R\} - U\{R\}|/q(|R|) \leq \sum_{k=1}^4 Z_{nk},$$

where with $\alpha_n = q^2(\frac{1}{n})$ and $\beta_n = \gamma \log n/n$ the rv's Z_{nk} are given by

$$(3.44) \quad Z_{n1} = \sup_{\beta_n \leq |R| \leq \alpha_n} |U_n\{R\}|/q(|R|),$$

$$(3.45) \quad Z_{n2} = \sup_{\alpha_n \leq |R| \leq \delta} |U_n\{R\}|/q(|R|),$$

$$(3.46) \quad Z_{n3} = \sup_{0 \leq |R| \leq \delta} |U\{R\}|/q(|R|),$$

$$(3.47) \quad Z_{n4} = \sup_{R \in \bar{R}} |U_n\{R\} - U\{R\}|/q(\delta).$$

It will be shown that for any $\varepsilon > 0$ and each $k = 1, 2, 3, 4$ there exist $\delta = \delta(\varepsilon)$ and $N = N(\varepsilon) \in \mathbb{N}$ such that

$$(3.48) \quad P(Z_{nk} \geq \varepsilon) \leq \varepsilon \text{ for } n \geq N.$$

For $k = 1$ the left side of (3.48) is bounded from above by

$$(3.49) \quad \begin{aligned} & P\left(\sup_{\beta_n \leq |R| \leq \alpha_n} |U_n\{R\}|/|R|^{\frac{1}{2}} \geq \varepsilon g(\alpha_n) (\log(1/\alpha_n))^{\frac{1}{2}}\right) \\ & \leq P\left(\sup_{\beta_n \leq |R| \leq \alpha_n} |U_n\{R\}|/|R|^{\frac{1}{2}} \geq \gamma_1 (\log n)^{\frac{1}{2}}\right) \end{aligned}$$

for $\gamma_1 \in (0, \infty)$ arbitrary and $n \geq N_1 = N_1(\gamma_1)$. Now Inequality 2.7 may be applied with $q(\cdot) = (\cdot)^{\frac{1}{2}}$ and $\varepsilon = \frac{1}{2}$. This yields the existence of $c_1, \dots, c_4 \in (0, \infty)$ such that the last expression of (3.49) is bounded from above by

$$(3.50) \quad \begin{aligned} & c_1 n (\log n)^{d-2} \exp(-c_2 \gamma_1^2 \log n \psi(c_3 \gamma_1)) \\ & \leq c_1 n (\log n)^{d-2} \exp(-c_4 \gamma_1 \log \gamma_1 \log n) \leq \varepsilon \end{aligned}$$

provided γ_1 and n are chosen sufficiently large.

To handle Z_{n2} we may again use Inequality 2.7. The integral in the resulting upper bound decreases to 0 as $n \rightarrow \infty$ and then $\delta \downarrow 0$, since $q \in Q_1$ implies

$$(3.51) \quad \int_0^1 \frac{(\log 1/\sigma)^{d-1}}{\sigma^2} \exp\left(\frac{-\lambda q^2(\sigma)}{\sigma}\right) d\sigma < \infty, \text{ for all } \lambda > 0, d \in \mathbf{N},$$

which can be seen by a slight modification of the proof of Proposition 3.1 in Shorack & Wellner (1982).

Using Theorem 2.1 in Orey and Pruitt (1973) we may treat Z_{n3} in the same way as Y_{n4} in the preceding section. We also have similarity between Z_{n4} and Y_{n5} using (3.39) instead of (1.8).

(\Rightarrow) For this half of the proof we refer to Csörgö, Csörgö, Horváth & Mason (1986, section 4.6) where the proof is given for the quantile process and the onedimensional empirical process. Their proof immediately carries over to the multivariate empirical process; the generalization of the results required in that paper can be found in Chapter 5, section 3, of this monograph and in Pyke (1972, p. 340) respectively.

□

We note in passing that the analogue for rectangles of Proposition 2.1 in O'Reilly (1974) can be obtained using some of the ideas in the proof of

Theorem 3.3: let $d \in \mathbb{N}$ and $q \in Q^*$. Then we have

$$(3.52) \quad \lim_{\delta \downarrow 0} \sup_{|R| \leq \delta} |U\{R\}|/q(|R|) = 0 \quad \text{a.s.}$$

if and only if $q \in Q_1$.

For any $\gamma \in (0, \infty)$ define $U_{n,\gamma}$, a process indexed by rectangles, by

$$(3.53) \quad U_{n,\gamma}\{R\} = U_n\{R\} 1_{[\gamma \log n/n, 1]}(|R|), \quad R \in \bar{R}.$$

Combining Theorem 3.3 and (3.52) yields

COROLLARY 3.3. Let $d \in \mathbb{N}$, $q \in Q^*$ and $\gamma \in (0, \infty)$ fixed. Then we have

$$(3.54) \quad \sup_{R \in \bar{R}} |U_{n,\gamma}\{R\} - U\{R\}|/q(|R|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if $q \in Q_1$.

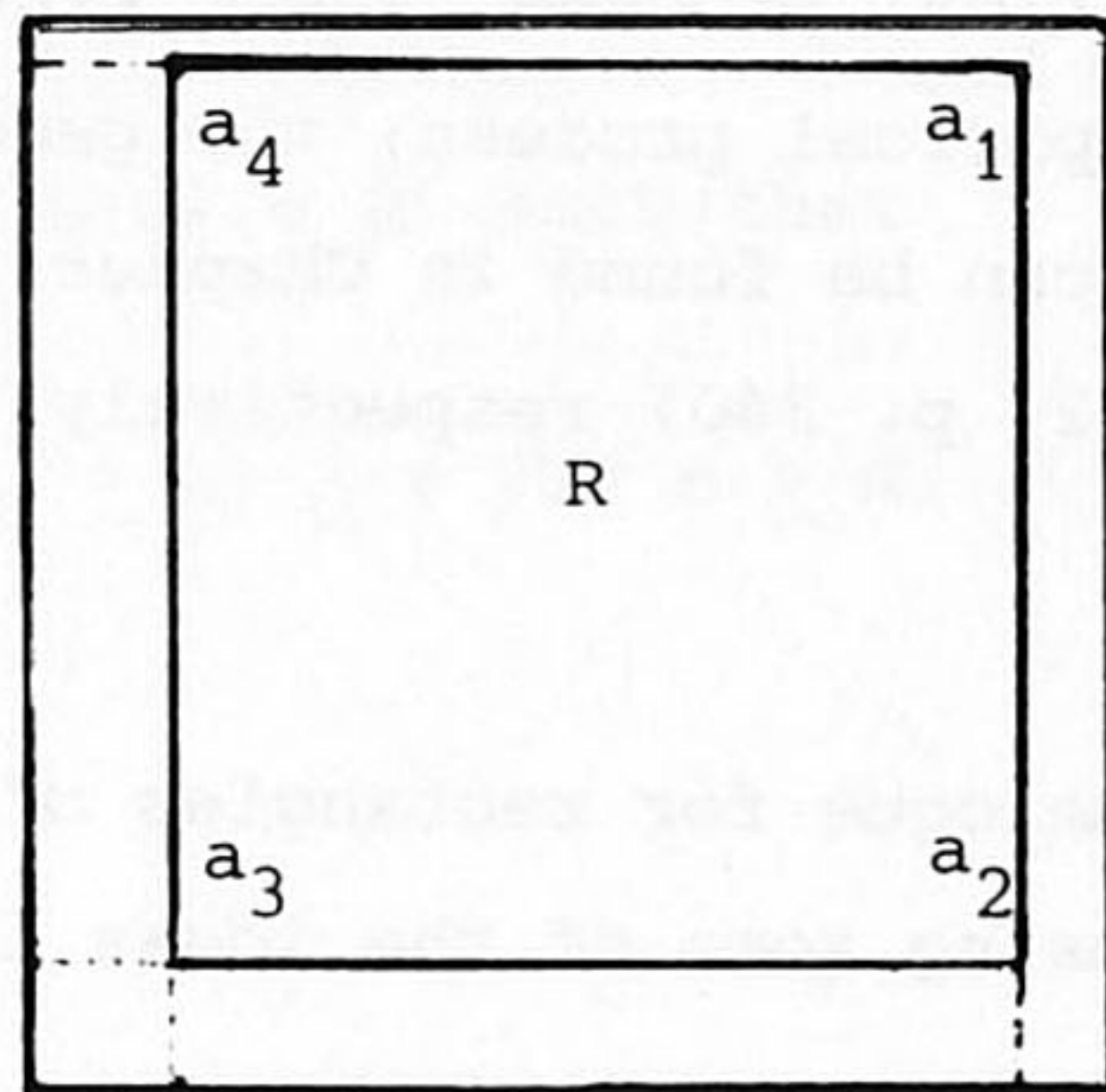
THEOREM 3.4. Let $d \in \mathbb{N}$ and $\tilde{q} \in Q^*$. Then we have

$$(3.55) \quad \sup_{R \in \bar{R}} |U_n\{R\} - U\{R\}|/\tilde{q}(1-|R|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if $\tilde{q} \in Q_0$.

PROOF. (\Leftarrow) For the sake of notational simplicity we restrict ourselves to the case $d = 2$. Without any difficulty the proof can be extended to arbitrary $d \in \mathbb{N}$. (See also the proof of Theorem 3.2.)

Note that it suffices to give this part of the proof with \bar{R} replaced by R in (3.55); cf. the proof of Lemma 3.1. So let us choose an arbitrary $R \in \bar{R}$ and denote its vertices with a_1, a_2, a_3, a_4 , starting at the right



upper vertex and moving clockwise. Note that $|a_1| \geq |R|$ and $|a_2|, |a_3|, |a_4| \leq 1 - |R|$. From the inequality

$$(3.56) \quad |U_n\{R\} - U\{R\}| / \tilde{q}(1 - |R|) \leq \sum_{i=1}^4 |U_n(a_i) - U(a_i)| / \tilde{q}(1 - |R|)$$

we see that we only have to deal with $\sup_{R \in \mathcal{R}} |U_n(a_i) - U(a_i)| / \tilde{q}(1 - |R|)$ for $i = 1, 2, 3, 4$. Using $\tilde{q}(1 - |R|) \geq \tilde{q}(1 - |a_1|)$ we can apply Theorem 3.2 to handle the case $i = 1$ and with the same technique as used in the proof of this theorem we can also treat the cases $i = 2, 3, 4$.

(\Rightarrow) Theorem 3.2 together with the observation that (3.55) implies (3.22) yields this part of the proof. □

Combination of Theorems 3.3 and 3.4 and Corollary 3.3 leads to

COROLLARY 3.4. *Let $d \in \mathbb{N}$, $q, \tilde{q} \in \mathcal{Q}^*$ and $\gamma \in (0, \infty)$ fixed. Then the following three statements are equivalent:*

$$(3.57) \quad \sup_{\gamma \log n / n \leq |R| \leq 1} \frac{|U_n\{R\} - U\{R\}|}{q(|R|)\tilde{q}(1 - |R|)} \xrightarrow{P} 0 \text{ as } n \rightarrow \infty,$$

$$(3.58) \quad \sup_{R \in \tilde{\mathcal{R}}} \frac{|U_{n,\gamma}\{R\} - U\{R\}|}{q(|R|)\tilde{q}(1 - |R|)} \xrightarrow{P} 0 \text{ as } n \rightarrow \infty,$$

$$(3.59) \quad q \in \mathcal{Q}_1 \text{ and } \tilde{q} \in \mathcal{Q}_0.$$

3. Discussion and bibliography

Chapter 3, apart from Corollary 3.2, is contained in Einmahl, Ruymgaart & Wellner (1984). Our theorems are also contained in Alexander (1982, 1984c, 1985), either implicitly or explicitly. It should be noted, however, that our approach is quite different from his.

Theorems 3.1, 3.2 and 3.4 are implicitly contained in Theorem 6.3 in Alexander (1982), which is proved, however, under the added assumption that

$$(3.60) \quad q(\cdot)/(\cdot) \text{ is non-increasing.}$$

Furthermore our Theorem 3.1 is explicitly given as his Corollary 6.10, but again under the additional assumption in (3.60). In Alexander (1984c), the aforementioned theorems are also obtained without assuming (3.60). In the onedimensional case Theorem 3.1 is established by Chibisov (1964) and O'Reilly (1974). For a discussion of Chibisov-O'Reilly weight functions and related topics, see Csörgö, Csörgö, Horváth & Mason (1983, section 2). Recently a new approximation of the onedimensional uniform empirical process has been obtained in Csörgö, Csörgö, Horváth & Mason (1986), which among other things yields a nice proof of the Chibisov-O'Reilly theorem. In this paper it is also shown that our Corollary 3.2 is not optimal: they prove, for $d = 1$, that

$$\sup_{t \in I} |U_n(t)| / q^*(t) q^*(1-t) \rightarrow_w \sup_{t \in I} |U(t)| / q^*(t) q^*(1-t)$$

with e.g. $q^*(\sigma) = (\sigma \log \log(1/\sigma))^{\frac{1}{2}}$, $\sigma \in [0,1]$. This q^* , however, is not an element of \mathcal{Q}_0 , and thus there is no weak convergence of the weighted processes themselves.

Theorem 3.3 is implicitly contained in Corollary 2.6 in Alexander (1985), where the if part is proved under the assumption in (3.60) and the only if part under the assumption that $q \in \mathcal{Q}$. For the if part of this theorem, see also Theorem 5.2 in Alexander (1982). Our truncation point $\gamma \log n/n$ in Theorem 3.3 is also chosen in Shorack & Wellner (1982, Theorem 1.2), where this theorem is proved for $d = 1$. Their proof of the only if part, however, is only correct for $q \in \mathcal{Q}$, and hence our theorem with $d = 1$ is an improvement of their theorem.

The choice of $\gamma \log n/n$ as truncation point is not as arbitrary as it seems to be at first sight. The reason for this choice is rather technical: $\{\gamma \log n/n\}_{n=1}^{\infty}$ is (roughly speaking) the smallest sequence for which the argument of ψ in (3.50) does not tend to ∞ as $n \rightarrow \infty$; for more details see Corollary 2.5 and the discussion above it in Alexander (1985). Finally note that in principle a corollary similar to Corollary 3.2 can be inferred from Corollary 3.4 by use of the point representation for rectangles. We did not present it due to the fact that the point representation of $U_{n,\gamma}$ is not an element of D_{2d} which causes some technical problems.

CHAPTER 4

STRONG LIMIT THEOREMS FOR SUPREMA OF
WEIGHTED MULTIVARIATE UNIFORM EMPIRICAL PROCESSES

This chapter will be devoted to strong limit theorems for suprema of the absolute values of weighted multivariate uniform empirical processes indexed by points. Here the weight functions q and \tilde{q} in (1.13) will be restricted to $q(\cdot) = \tilde{q}(\cdot) = (\cdot)^\alpha$, $0 \leq \alpha \leq 1$. In section 1 the suprema will be taken over the whole unit square, in section 2 over its middle ($= \{t \in I^d : |t| \text{ bounded away from } 0\}$) and in section 3 over its tail ($= \{t \in I^d : |t| \text{ close to } 0\}$). Some refinements of the results can and will be given for the onedimensional case in sections 1 and 3. These refinements have no natural generalization to arbitrary dimension since they depend on the concept of an order statistic. It turns out that the class of rectangles is not a very suitable indexing class for the investigations in this chapter. Therefore we only briefly consider this indexing class in section 4.

1. Suprema over the whole unit square: proper standardization is impossible

We begin this section with a well-known result for order statistics, a version of the Borel - Cantelli lemma and an inequality. All these results turn out to be extremely useful in this chapter. For $d = 1$ and any integer $1 \leq k \leq n$ let $X_{k:n}$ denote the k -th order statistic of X_1, X_2, \dots, X_n .

FACT 4.1. (Mori (1976); see also Kiefer (1972) and Geffroy (1958/1959).)
For any fixed positive integer k and sequence of positive constants $\{b_n\}_{n=1}^\infty$ with $b_n \downarrow$ we have

$$(4.1) \quad \sum_{n=1}^{\infty} n^{k-1} b_n^k = \infty \text{ implies } P(X_{k:n} \leq b_n \text{ i.o.}) = 1.$$

For $k = 1$ statement (4.1) holds true without the condition $b_n \downarrow$.

FACT 4.2. (Barndorff-Nielsen (1961).) Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of events in some probability space. Then $\sum_{n=2}^{\infty} P(A_n^c | A_{n-1}^c) < \infty$ and $\lim_{n \rightarrow \infty} P(A_n) = 0$ imply $P(A_n \text{ i.o.}) = 0$.

INEQUALITY 4.1. For every $d \in \mathbb{N}$ and integer m with $1 \leq m \leq n$ and $0 < \alpha \leq 1$ we have

$$(4.2) \quad P\left(\sup_{|t| \leq \alpha} nF_n(t) \geq m\right) \leq c_1 \binom{n}{m} (c_2 \alpha)^m (1 \vee \log(1/\alpha))^{d-1},$$

where $c_1 = c_1(d)$, $c_2 = c_2(d)$ and $c_1, c_2 \in [1, \infty)$. Of course $c_1(1)$ and $c_2(1)$ can both be taken equal to 1.

PROOF. Using the method of proof of Inequality 2.6 we obtain, with $P(\frac{1}{2}; 0, \alpha)$ defined by (2.24),

$$\begin{aligned} (4.3) \quad & P\left(\sup_{|t| \leq \alpha} nF_n(t) \geq m\right) \\ & \leq P\left(\sup_{R(a,b) \in P(\frac{1}{2}; 0, \alpha)} \sup_{t \in R(a,b)} nF_n(t) \geq m\right) \\ & \leq \sum_{R(a,b) \in P(\frac{1}{2}; 0, \alpha)} P\left(\sup_{t \in R(a,b)} nF_n(t) \geq m\right) \\ & \leq \sum_{R(a,b) \in P(\frac{1}{2}; 0, \alpha)} P(nF_n(b) \geq m) \\ & \leq \binom{n}{m} \sum_{R(a,b) \in P(\frac{1}{2}; 0, \alpha)} |b|^m \\ & \leq \binom{n}{m} 2^{dm} 2^d \int_{|t| \leq 2^d \alpha} |t|^{m-1} |dt| \\ & \leq \binom{n}{m} 2^{dm} 2^d \int_0^{2^d \alpha} (\log(1/s))^{d-1} s^{m-1} ds. \end{aligned}$$

Elementary analysis shows

$$(4.4) \quad \int_0^{2^d \alpha} (\log(1/s))^{d-1} s^{m-1} ds \leq c(d) (2^d \alpha)^m (1 \vee \log(1/\alpha))^{d-1},$$

which completes the proof.

□

We are now prepared to state and prove the two theorems of this section and some corollaries. The second theorem is a refinement of the first one for the case $d = 1$. We first state the results and afterwards the proofs.

Let us define

$$(4.5) \quad V_{n,v} = \sup_{t \in I^d} \frac{|U_{n,v}(t)|}{(|t|(1-|t|))^{1-v}}, \quad 0 \leq v \leq \frac{1}{2},$$

and for $d = 1$ and fixed positive integer k

$$(4.6) \quad V_{n,v}^{(k)} = \sup_{X_{k:n-1} < t < X_{n-k+1:n}} \frac{|U_{n,v}(t)|}{(t(1-t))^{1-v}}, \quad 0 \leq v \leq \frac{1}{2},$$

with $U_{n,v}$ as defined in (2.51). Also let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive constants.

THEOREM 4.1. Let $d \in \mathbb{N}$ and $0 \leq v \leq \frac{1}{2}$.

(I) If $\sum_{n=1}^{\infty} a_n (\log(1/a_n))^{d-1} = \infty$, then

$$(4.7) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-v} V_{n,v} = \infty \quad \text{a.s.}$$

(II) If $\sum_{n=1}^{\infty} a_n (\log(1/a_n))^{d-1} < \infty$ and $na_n \downarrow$, then

$$(4.8) \quad \lim_{n \rightarrow \infty} (na_n)^{1-v} V_{n,v} = 0 \quad \text{a.s.}$$

COROLLARY 4.1. There exists no sequence of positive constants $\{a_n\}_{n=1}^{\infty}$ such that $na_n \downarrow$ and

$$(4.9) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-v} V_{n,v} = 1 \quad \text{a.s.}$$

COROLLARY 4.2.

$$(4.10) \quad \limsup_{n \rightarrow \infty} \frac{\log V_{n,v}}{\log \log n} = (1-v)d \quad \text{a.s.}$$

COROLLARY 4.3. For all $\beta < 0$

$$(4.11) \quad \lim_{n \rightarrow \infty} n^\beta V_{n,v} = 0 \quad \text{a.s.}$$

THEOREM 4.2. Let $d = 1$, $k \in \mathbb{N}$ and $0 \leq v \leq \frac{1}{2}$.

(I) If $\sum_{n=1}^{\infty} n^{k-1} a_n^k = \infty$ and $a_n \downarrow$, then

$$(4.12) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-v} V_{n,v}^{(k)} = \infty \quad \text{a.s.}$$

(II) If $\sum_{n=1}^{\infty} n^{k-1} a_n^k < \infty$ and $na_n \downarrow$, then

$$(4.13) \quad \lim_{n \rightarrow \infty} (na_n)^{1-v} V_{n,v}^{(k)} = 0 \quad \text{a.s.}$$

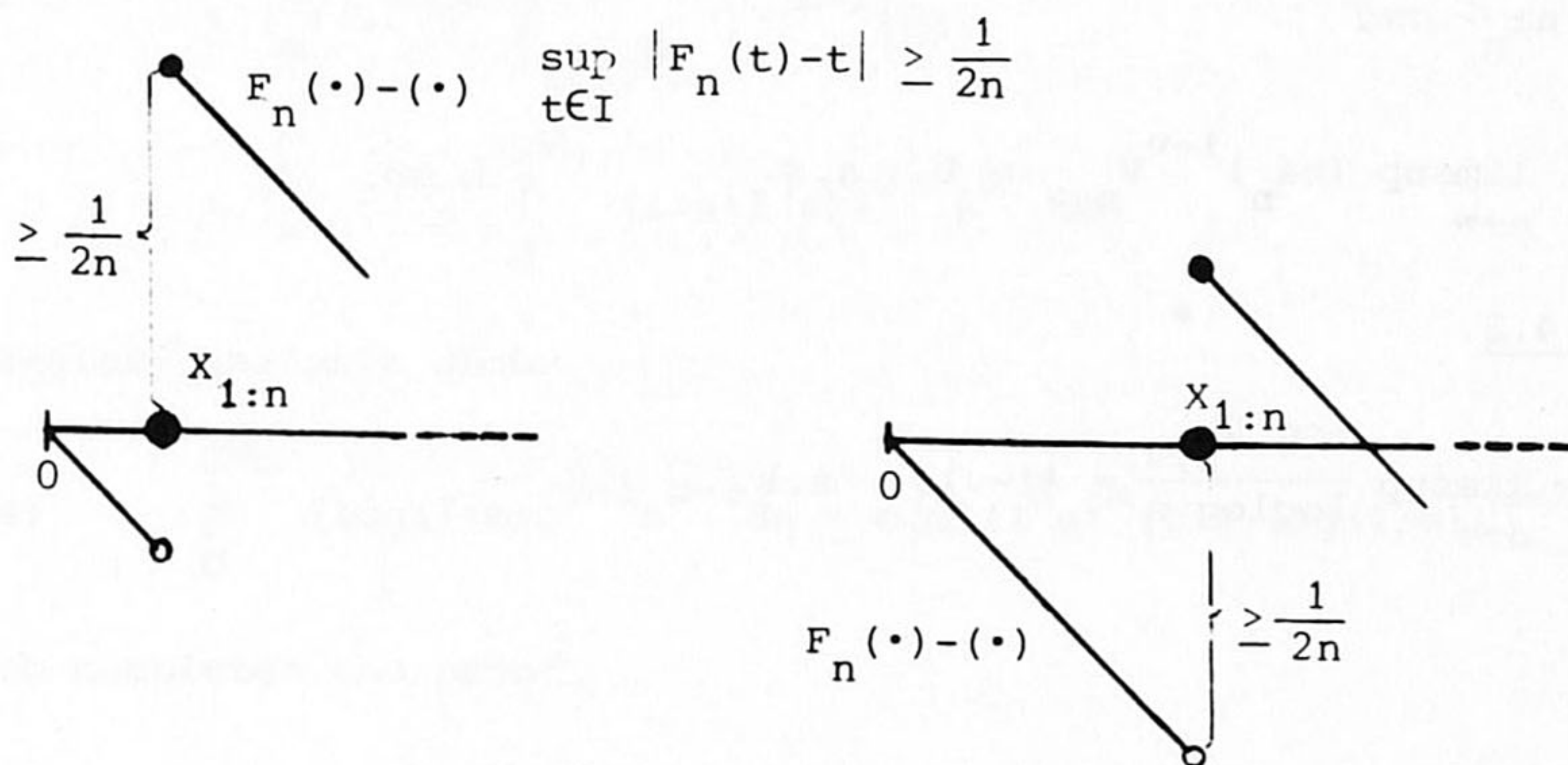
COROLLARY 4.4.

$$(4.14) \quad \limsup_{n \rightarrow \infty} \frac{\log V_{n,v}^{(k)}}{\log \log n} = \frac{1-v}{k} \quad \text{a.s.}$$

PROOF OF THEOREM 4.1. Part I. It is a consequence of Fact 4.1 with $k = 1$ that $\sum a_n \log(1/a_n)^{d-1} = \infty$ implies $P(|X|_{1:n} < \varepsilon a_n \text{ i.o.}) = 1$ for any $\varepsilon > 0$, with $|X|_{1:n} = \min\{|X_1|, \dots, |X_n|\}$; see the remark below (3.13). It can be easily seen that

$$(4.15) \quad (na_n)^{1-v} V_{n,v} \geq \frac{(na_n)^{1-v} n^{v-1}}{2(|X|_{1:n} (1-|X|_{1:n}))^{1-v}}$$

$$d = 1$$



Hence we have

$$(4.16) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-\nu} V_{n,\nu} \geq \frac{1}{2\varepsilon^{1-\nu}} \quad \text{a.s.}$$

Letting $\varepsilon \downarrow 0$ proves this part of the theorem.

Part II. We restrict ourselves without loss of generality to sequences $\{a_n\}_{n=1}^{\infty}$ with $n^{-2} \leq a_n \leq n^{-1}$. Using $a_n \leq n^{-1}$ we see that the assumptions of part II imply $\sum a_n (\log n)^{d-1} < \infty$, which in turn in combination with $na_n \downarrow$ implies $na_n (\log n)^d \rightarrow 0$ as $n \rightarrow \infty$. Define for either choice of sign

$$(4.17) \quad V_n^{\pm} = \sup_{0 < |t| \leq b_n} \frac{\pm(F_n(t) - |t|)}{|t|^{1-\nu}},$$

with $b_n = \log n/n$ and let us first show that

$$(4.18) \quad \lim_{n \rightarrow \infty} na_n^{1-\nu} V_n^+ = 0 \quad \text{a.s.}$$

It suffices to prove that $\limsup_{n \rightarrow \infty} na_n^{1-\nu} V_n^+ \leq 1$ a.s.

Define the following events:

$$(4.19) \quad A_n = \{V_n^+ \geq \frac{1}{na_n^{1-\nu}}\}; \quad C_n = A_n^c A_{n-1}^c.$$

According to Fact 4.2 we need to prove that $\sum PC_n < \infty$ and $\lim_{n \rightarrow \infty} PA_n = 0$.

Let $y_{i,n}$ be the solution of the equation

$$(4.20) \quad f_n(\sigma) = i, \text{ for } i = 0, 1, 2, \dots, i_n = [f_n(b_n)] + 1,$$

where

$$(4.21) \quad f_n(\sigma) = n\sigma + (\sigma/a_n)^{1-\nu}, \quad \sigma \in [0, \infty).$$

Note that $i_n = 1 + [\log n + ((\log n)/(na_n))^{1-\nu}]$ and hence, since $na_n \log n \rightarrow 0$ as $n \rightarrow \infty$, we have $i_n \leq 3((\log n)/(na_n))^{1-\nu}$. Moreover, let $x_i = x_{i,n} = y_{i,n} \wedge b_n$ and observe that for large n , f_n is increasing on $[0, \infty)$, $f_{n-1} \leq f_n$, $b_{n-1} \geq b_n$, $x_{i_n} = b_n$ and $1/(2n)^2 \leq x_i \leq i^{1/(1-\nu)} a_n$. Hence we have for large n

$$\begin{aligned}
(4.22) \quad C_n &= \left\{ \sup_{0 < |t| \leq b_n} \frac{F_n(t) - |t|}{|t|^{1-\nu}} \geq \frac{1}{na_n^{1-\nu}}; \sup_{0 < |t| \leq b_{n-1}} \frac{F_{n-1}(t) - |t|}{|t|^{1-\nu}} < \frac{1}{(n-1)a_{n-1}^{1-\nu}} \right\} \\
&\subset \{ \exists_{t: 0 < |t| \leq b_n} nF_n(t) \geq f_n(|t|); \forall_{t: 0 < |t| \leq b_n} (n-1)F_{n-1}(t) < f_n(|t|) \} \\
&\subset \bigcup_{i=1}^{i_n} \{ \exists_{t: x_{i-1} < |t| \leq x_i} nF_n(t) = i; \forall_{t: x_{i-1} < |t| \leq x_i} (n-1)F_{n-1}(t) \leq i-1 \} \\
&\subset \bigcup_{i=1}^{i_n} B_{n,i},
\end{aligned}$$

where

$$(4.23) \quad B_{n,i} = \left\{ \sup_{|t| \leq x_i} (n-1)F_{n-1}(t) \geq i-1; |x_n| \leq x_i \right\}.$$

We are now going to compute upper bounds for $PB_{n,i}$ ($i = 1, 2, \dots, i_n$) and hence for PC_n . In this computation c_1, c_2, \dots, c_9 are finite numbers only depending on d and ν . We first consider $i = 1$. Then we have by Inequality 4.1 and the remarks above (4.22) for large n

$$(4.24) \quad PB_{n,1} \leq c_1 c_2 x_1 (\log(1/x_1))^{d-1} \leq c_3 a_n (\log n)^{d-1}.$$

For $2 \leq i \leq i_n$ we have similarly for large n

$$(4.25) \quad PB_{n,i} \leq c_1^2 \binom{n-1}{i-1} (c_2 x_i)^i (\log(1/x_i))^{2d-2} \leq \frac{c_4}{n} \frac{(c_2 n x_i)^i}{(i-1)!} (\log n)^{2d-2}.$$

Using the Stirling formula we obtain

$$(4.26) \quad 1/(m!) \leq (e/m)^m \quad \forall_{m \in \mathbb{N}}.$$

Combination of (4.25), (4.26) and the remarks above (4.22) yields

$$\begin{aligned}
(4.27) \quad PB_{n,i} &\leq c_5 n x_i^{2i-1} \left(\frac{c_6 n x_i}{i} \right)^{i-2} (\log n)^{2d-2} \\
&\leq c_5 n_i^{\frac{1+\nu}{1-\nu}} a_n^2 (c_6 n_i^{\frac{1-\nu}{1-\nu}} a_n)^{i-2} (\log n)^{2d-2} \\
&\leq c_5 i^3 n a_n^2 (3c_6 (n a_n \log n)^{\frac{1}{2}})^{i-2} (\log n)^{2d-2}.
\end{aligned}$$

Since $na_n \log n \rightarrow 0$ as $n \rightarrow \infty$, this last expression is less than or equal to

$$(4.28) \quad c_7 na_n^2 \left(\frac{1}{2}\right)^{i-2} (\log n)^{2d-2}.$$

Combining (4.22), (4.24), (4.27) and (4.28) yields for large n

$$(4.29) \quad \begin{aligned} PC_n &\leq c_3 a_n (\log n)^{d-1} + c_7 na_n^2 (\log n)^{2d-2} \sum_{i=2}^{\infty} \left(\frac{1}{2}\right)^{i-2} \\ &\leq c_3 a_n (\log n)^{d-1} (1 + c_8 na_n (\log n)^{d-1}) \leq c_9 a_n (\log n)^{d-1}. \end{aligned}$$

Now we have, since $\sum a_n (\log n)^{d-1} < \infty$, that $\sum PC_n < \infty$.

For the proof of $PA_n \rightarrow 0$ as $n \rightarrow \infty$, we need

$$(4.30) \quad A_{n,i} = \left\{ \sup_{|t| \leq x_i} nF_n(t) \geq i \right\}.$$

Since $x_{i_n} = b_n$ we have

$$(4.31) \quad A_n \subset \bigcup_{i=1}^{i_n} A_{n,i}.$$

Using Inequality 4.1 we have for $1 < i \leq i_n$ by a similar computation as for the $B_{n,i}$

$$(4.32) \quad PA_{n,i} \leq c(d, \nu) \left(\frac{1}{2}\right)^{i-1} na_n (\log n)^{d-1},$$

which in combination with (4.31) yields for large n

$$(4.33) \quad PA_n \leq 2c(d, \nu) na_n (\log n)^{d-1} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Our next step is investigating the almost sure behaviour of V_n^- as defined in (4.17). We immediately see that

$$(4.33) \quad na_n^{1-\nu} V_n^- \leq na_n^{1-\nu} \sup_{0 < |t| \leq b_n} |t|^\nu = (na_n \log n)^{1-\nu} (\log n)^{2\nu-1},$$

which yields that

$$(4.34) \quad \lim_{n \rightarrow \infty} na_n^{1-\nu} V_n^- = 0 \quad \text{a.s.}$$

Now consider

$$(4.35) \quad V_n^* = \sup_{b_n < |t| \leq 1} \frac{|F_n(t) - |t||}{|t|^{1-v}}.$$

From Theorem 4.3 in the next section it easily follows that

$$(4.36) \quad \limsup_{n \rightarrow \infty} \sup_{b_n < |t| \leq 1} \frac{|U_{n,v}(t)|}{(\log \log n)^v |t|^{1-v}} \leq (2(d+1))^{\frac{1}{2}} \quad \text{a.s.}$$

Using $na_n \log n \rightarrow 0$ as $n \rightarrow \infty$, (4.36) yields that

$$(4.37) \quad \lim_{n \rightarrow \infty} na_n^{1-v} V_n^* = 0 \quad \text{a.s.}$$

Combining (4.18), (4.34) and (4.37) yields

$$(4.38) \quad \lim_{n \rightarrow \infty} na_n^{1-v} \sup_{t \in I^d} \frac{|F_n(t) - |t||}{|t|^{1-v}} = 0 \quad \text{a.s.}$$

Noting that $0 < y \leq \frac{1}{2}$ implies $1 < (1-y)^{v-1} \leq 2$, we see that it remains to prove for $0 \leq v \leq \frac{1}{2}$

$$(4.39) \quad \lim_{n \rightarrow \infty} (na_n)^{1-v} \sup_{\frac{1}{2} < |t| < 1} \frac{|U_{n,v}(t)|}{(1-|t|)^{1-v}} = 0 \quad \text{a.s.}$$

With the same approach as in the proof of Theorem 3.2 we can prove, using Inequality 2.6 or a result in Alexander (1982, Corollary 6.2) that "large d -dimensional points" behave as "small (or large) 1-dimensional points", i.e.

$$(4.40) \quad \sum_{n=1}^{\infty} a_n' < \infty \text{ and } na_n' \downarrow \text{ imply}$$

$$\lim_{n \rightarrow \infty} (na_n')^{1-v} \sup_{t \in I^d} \frac{|U_{n,v}(t)|}{(1-|t|)^{1-v}} = 0 \quad \text{a.s.}$$

We omit the proof of (4.40), because it is straightforward though tedious. □

PROOF OF COROLLARY 4.2. Applying Theorem 4.1 with $a_n = (n(\log n)^d)^{-1}$ and with $a_n = (n(\log n)^{d+\epsilon})^{-1}$, $\epsilon > 0$, gives the desired result. □

PROOF OF THEOREM 4.2. Part I. Since $\sum n^{k-1} a_n^k = \infty$ and $a_n \downarrow$ it follows from Fact 4.1 that for every $\varepsilon > 0$

$$(4.41) \quad P(X_{k:n} \leq \varepsilon a_n \wedge k/(2n) \text{ i.o.}) = 1.$$

It is easy to see that

$$(4.42) \quad (na_n)^{1-\nu} V_{n,\nu}^{(k)} \geq (na_n^{1-\nu} (\frac{k}{n} - X_{k:n})) / X_{k:n}^{1-\nu},$$

hence (4.41) implies for every $\varepsilon > 0$

$$(4.43) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-\nu} V_{n,\nu}^{(k)} \geq k/(2\varepsilon^{1-\nu}) \quad \text{a.s.}$$

Letting $\varepsilon \downarrow 0$ proves this part of the theorem.

Part II. The proof is similar to the proof of the corresponding part of Theorem 4.1. We restrict ourselves without loss of generality to sequences $\{a_n\}_{n=1}^{\infty}$ with $a_n \geq (n(\log n)^2)^{-1}$. By the assumptions of part II we also have $na_n(\log n)^{1/k} \rightarrow 0$ as $n \rightarrow \infty$. Define for either choice of sign

$$(4.44) \quad W_n^{\pm} = \begin{cases} \sup_{X_{k:n} \leq t \leq b_n^*} \frac{\pm(F_n(t) - t)}{t^{1-\nu}} & \text{if } X_{k:n} \leq b_n^* \\ 0 & \text{otherwise,} \end{cases}$$

with $b_n^* = (\log n)^{1/k} / n$ and let us first show that

$$(4.45) \quad \limsup_{n \rightarrow \infty} na_n^{1-\nu} W_n^+ \leq 0 \quad \text{a.s.}$$

It is sufficient to prove $\limsup_{n \rightarrow \infty} na_n^{1-\nu} W_n^+ \leq 1$ a.s.

According to Fact 4.2 we need to prove $\sum PC_n^* < \infty$ and $\lim_{n \rightarrow \infty} PA_n^* = 0$, where now

$$(4.46) \quad A_n^* = \{W_n^+ \geq \frac{1}{na_n^{1-\nu}}\} ; C_n^* = A_n^* A_{n-1}^{*c}.$$

Let $f_n(\sigma)$ and $y_{i,n}$ be the quantities defined in the proof of Theorem 4.1 and set $x_i^* = x_{i,n}^* = y_{i,n} \wedge b_n^*$ and $i_n^* = [f_n(b_n^*)] + 1$. We can make similar

observations as between (4.21) and (4.22), in particular we have $i_n^* \leq 3((\log n)^{1/k} / (na_n))^{1-\nu}$ for large n . To establish (4.45) we also require the following statements, which will be proved for large n :

$$(4.47) \quad \{W_{n-1}^+ < \frac{1}{(n-1)a_{n-1}^{1-\nu}}\} \subset \{\forall_{t: x_{k-1,n}^* < t \leq b_n^*} (n-1)F_{n-1}(t) < f_n(t)\},$$

$$(4.48) \quad \{W_n^+ \geq \frac{1}{na_n^{1-\nu}}\} \subset \bigcup_{i=k}^{i_n^*} \{\exists_{t: x_{i-1,n}^* < t \leq x_{i,n}^*} nF_n(t) \geq f_n(t)\},$$

$$(4.49) \quad \{W_{n-1}^+ < \frac{1}{(n-1)a_{n-1}^{1-\nu}}; W_n^+ \geq \frac{1}{na_n^{1-\nu}}\} \subset \bigcup_{i=k}^{i_n^*} \{(n-1)F_{n-1}(x_{i,n}^*) \geq i-1; \\ x_n \leq x_{i,n}^*\}.$$

We first prove (4.47). Let $\omega \in \{W_{n-1}^+ < ((n-1)a_{n-1}^{1-\nu})^{-1}\}$. We have to distinguish between the cases $X_{k:n-1} > b_n^*$ and $X_{k:n-1} \leq b_n^*$. In the first case we have $(n-1)F_{n-1}(t) \leq k-1$ for all $0 < t \leq b_n^*$, which implies that $(n-1)F_{n-1}(t) < nt + (t/a_n)^{1-\nu}$ for all $x_{k-1,n}^* < t \leq b_n^*$. In the second case we have $(n-1)F_{n-1}(t) < (n-1)t + (t/a_{n-1})^{1-\nu}$ for all $X_{k:n-1} \leq t \leq b_n^*$, which since $na_n \downarrow$ implies

$$(4.50) \quad (n-1)F_{n-1}(t) < nt + (t/a_n)^{1-\nu} \text{ for } X_{k:n-1} \leq t \leq b_n^*.$$

Notice that if $X_{k:n-1} \leq y_{k-1,n}$, then $(n-1)F_{n-1}(y_{k-1,n}) \geq k$, which contradicts (4.50). Thus if (4.50) holds we must have

$$(4.51) \quad X_{k:n-1} > y_{k-1,n}.$$

Now (4.51) implies that

$$(4.52) \quad (n-1)F_{n-1}(t) < nt + (t/a_n)^{1-\nu} \text{ for } y_{k-1,n} < t < X_{k:n-1}.$$

Combining (4.50) and (4.52) and replacing $y_{k-1,n}$ by $x_{k-1,n}^*$ yields

$$(4.53) \quad \omega \in \{\forall_{t: x_{k-1,n}^* < t \leq b_n^*} (n-1)F_{n-1}(t) < nt + (t/a_n)^{1-\nu}\}.$$

This completes the proof of (4.47).

Now we will prove (4.48). Let $\omega \in \{W_n^+ \geq (na_n^{1-\nu})^{-1}\}$. We must have $X_{k:n} \leq b_n^*$. Since $x_{i,n}^* = b_n^*$ there exists $1 \leq i \leq i_n^*$ such that

$$(4.54) \quad X_{k:n} \in (x_{i-1,n}^*, x_{i,n}^*].$$

If (4.54) holds with $i < k$ then $nF_n(x_{k,n}^*) \geq k$, which implies that ω is an element of the set on the right side of (4.48). If (4.54) holds with $i \geq k$, then there must exist a t with $y_{k-1,n} = x_{k-1,n}^* < X_{k:n} \leq t \leq b_n^*$ and $nF_n(t) \geq nt + (t/a_n)^{1-\nu}$, which also implies that ω is an element of the set on the right side of (4.48). This completes the proof of (4.48).

Finally we prove (4.49). Observe that for any $k \leq i \leq i_n^*$

$$(4.55) \quad \begin{aligned} & \{ \forall_{t: x_{k-1,n}^* < t \leq b_n^*} (n-1)F_{n-1}(t) < f_n(t); \exists_{t: x_{i-1,n}^* < t \leq x_{i,n}^*} nF_n(t) \geq f_n(t) \} \\ & \subset \{ \exists_{t: x_{i-1,n}^* < t \leq x_{i,n}^*} nF_n(t) \geq i; (n-1)F_{n-1}(x_{i,n}^*) < i \} \\ & \subset \{ (n-1)F_{n-1}(x_{i,n}^*) \geq i-1; X_n \leq x_{i,n}^* \}. \end{aligned}$$

Now (4.49) follows from (4.47) and (4.48).

We are now ready to show that $\sum PC_n^* < \infty$. Using (4.49) and some of the ideas of the proof of Theorem 4.1 we have for large n

$$(4.56) \quad \begin{aligned} PC_n^* & \leq \sum_{i=k}^{i_n^*} P(X_n \leq x_i^*) P((n-1)F_{n-1}(x_i^*) \geq i-1) \\ & \leq \sum_{i=k}^{i_n^*} \binom{n-1}{i-1} (x_i^*)^i \leq n^{k-1} a_n^k \sum_{i=k}^{i_n^*} e^{k i^{kv/(1-\nu)+1}} (e n a_n^{v/(1-\nu)})^{i-k} \\ & \leq n^{k-1} a_n^k \sum_{i=k}^{i_n^*} e^{k i^{kv/(1-\nu)+1}} (3e (n a_n (\log n)^{1/k})^{1/2})^{i-k} \\ & \leq n^{k-1} a_n^k \sum_{i=k}^{\infty} e^{k i^{k+1} (\frac{1}{2})} i^{-k} = c(k) n^{k-1} a_n^k. \end{aligned}$$

Hence we have, since $\sum n^{k-1} a_n^k < \infty$, that $\sum PC_n^* < \infty$.

Now we show that $\lim_{n \rightarrow \infty} PA_n^* = 0$. Observe that by (4.48) for large n

$$(4.57) \quad PA_n^* \leq \sum_{i=k}^{i_n^*} P(nF_n(x_i^*) \geq i) \leq \sum_{i=k}^{i_n^*} \binom{n}{i} (x_i^*)^i,$$

which by the same computation as in (4.56) is less than or equal to $c(k)n^k a_n^k$ for large n . Since $\lim_{n \rightarrow \infty} na_n = 0$ we have $\lim_{n \rightarrow \infty} PA_n^* = 0$ and hence (4.45).

In the same way as in the proof of Theorem 4.1 it can be shown that

$$(4.58) \quad \limsup_{n \rightarrow \infty} na_n^{1-\nu} W_n^- \leq 0 \quad \text{a.s.}$$

and that

$$(4.59) \quad \lim_{n \rightarrow \infty} na_n^{1-\nu} \sup_{b_n^* < t \leq 1} \frac{|F_n(t) - t|}{t^{1-\nu}} = 0 \quad \text{a.s.,}$$

hence we have

$$(4.60) \quad \lim_{n \rightarrow \infty} na_n^{1-\nu} \sup_{X_{k:n} < t < 1} \frac{|F_n(t) - t|}{t^{1-\nu}} = 0 \quad \text{a.s.}$$

Now by symmetry considerations (4.13) almost immediately follows from (4.60).

□

PROOF OF COROLLARY 4.4. Applying Theorem 4.2 with $a_n = (n(\log n)^{1/k})^{-1}$ and with $a_n = (n(\log n)^{1/k+\varepsilon})^{-1}$, $\varepsilon > 0$, gives the desired result.

□

2. Suprema over the middle: a law of the iterated logarithm

In this section let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive constants such that

$$(4.61) \quad \frac{na_n}{\log \log n} \rightarrow c \in (0, \infty) \text{ as } n \rightarrow \infty.$$

For any $\sigma > 0$ define β_σ by

$$(4.62) \quad \beta_\sigma (\log \beta_\sigma - 1) + 1 = \sigma^{-1} \text{ and } \beta_\sigma > 1.$$

We now state first a useful inequality for the multinomial distribution and then the theorem of this section and some (results which are almost) corollaries.

FACT 4.3. (Mallows (1968).) If $\langle N_1, \dots, N_m \rangle$, $m \in \mathbb{N}$, is multinomially distributed with parameters n and p_1, \dots, p_m , where $n \in \mathbb{N}$ and p_1, \dots, p_m are non-negative with $\sum_{i=1}^m p_i = 1$, then we have for any $\lambda_1, \dots, \lambda_m$

$$P(N_1 \leq \lambda_1; \dots; N_m \leq \lambda_m) \leq \prod_{i=1}^m P(N_i \leq \lambda_i).$$

THEOREM 4.3. Let $d \in \mathbb{N}$, $0 \leq v \leq \frac{1}{2}$ and $c \in (0, \infty)$. Then we have

$$(4.63) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |t| \leq 1} \frac{|U_{n,v}(t)|}{(\log \log n)^v |t|^{1-v}} = c^v (\beta_{c/d}^{-1})^v (2(d+1))^{\frac{1}{2}} 1_{\{\frac{1}{2}\}}(v) \text{ a.s.}$$

COROLLARY 4.5. For every fixed $\alpha \in (0, 1)$ we have

$$(4.64) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |t| \leq \alpha} \frac{|U_{n,v}(t)|}{(\log \log n)^v (|t|(1-|t|))^{1-v}} \\ = c^v (\beta_{c/d}^{-1})^v (2(d+1))^{\frac{1}{2}} 1_{\{\frac{1}{2}\}}(v) \text{ a.s.}$$

COROLLARY 4.6.

$$(4.65) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |t| \leq 1} |F_n(t)/|t| - 1| = \beta_{c/d}^{-1} \text{ a.s.,}$$

$$(4.66) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |t| \leq 1} F_n(t)/|t| = \beta_{c/d} \text{ a.s.}$$

Without loss of generality we restrict ourselves in the proofs of Theorem 4.3 and its corollaries to sequences $\{a_n\}_{n=1}^{\infty}$ with $a_n = cn^{-1} \log \log n$.

PROOF OF THEOREM 4.3. The case $v = \frac{1}{2}$ is due to Alexander (1984b, Corollary 3.5).

So assume $0 \leq v < \frac{1}{2}$. (The fact that we do not give a proof for $v = \frac{1}{2}$ is not imposed by the limitations of our approach. The case $v = \frac{1}{2}$, however, behaves somewhat differently with regard to the other values of v , so that in some parts of the proof we should have to distinguish between the cases $v = \frac{1}{2}$ and $0 \leq v < \frac{1}{2}$.)

We first consider the upper bound. We split the interval $[a_n, 1]$ in three subintervals: $[a_n, n^{-1} \log n]$, $[n^{-1} \log n, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ and prove first, with $A = c^v (\beta_{c/d}^{-1})^v$,

$$(4.67) \quad \limsup_{n \rightarrow \infty} \sup_{a_{n-1} \leq |t| \leq n^{-1} \log n} \frac{|U_{n,v}(t)|}{(\log \log n)^v |t|^{1-v}} \leq A \quad \text{a.s.}$$

By the Borel-Cantelli lemma and Inequality 2.10 it suffices to show that for every small positive ε we have $\sum PD_k < \infty$, where

$$(4.68) \quad D_k = \left\{ \sup_{a_{n_{k+1}} \leq |t| \leq n_k^{-1} \log n_k} \frac{|U_{n_{k+1},v}(t)|}{|t|^{1-v}} \geq (1+\varepsilon)(1-\varepsilon)^{-2} A (\log \log n_k)^v \right\}$$

and $n_k = [(1+\varepsilon/12)^k]$. (Note that when $v = 0$, Inequality 2.10 is not applicable for $\lambda \leq 1$; a slightly modified version of the inequality, however, does the work in this case.)

We now use Inequality 2.8, which yields

$$(4.69) \quad PD_k \leq C \int_{(1-\varepsilon)a_{n_{k+1}}}^{((1-\varepsilon)n_k)^{-1} \log n_k} \frac{(\log(1/\sigma))^{d-1}}{\sigma} d\sigma \cdot \exp(-\frac{1}{2}(1+\varepsilon)^2 (1-\varepsilon)^{-2} A^2 c^{1-2v} \log \log n_k \psi((1+\varepsilon)(1-\varepsilon)^{-2} A c^{-v})).$$

Using $\sigma^d - \tau^d = (\sigma - \tau) \sum_{\ell=0}^{d-1} \sigma^\ell \tau^{d-\ell-1}$ we see that for large k

$$(4.70) \quad \int_{(1-\varepsilon)a_{n_{k+1}}}^{((1-\varepsilon)n_k)^{-1} \log n_k} \frac{(\log(1/\sigma))^{d-1}}{\sigma} d\sigma = d^{-1} \left\{ \left(\log \left(\frac{n_{k+1}}{(1-\varepsilon)c \log \log n_{k+1}} \right) \right)^d - \left(\log \left(\frac{(1-\varepsilon)n_k}{\log n_k} \right) \right)^d \right\} \leq (\log n_k)^{d-1} \log \log n_k.$$

From the fourth property of ψ in (2.15) and from (4.62) combined with (2.14) it follows that the exponential in (4.69) is less than or equal to

$$(4.71) \quad \exp(-(1+\varepsilon)d \log \log n_k).$$

Combining (4.69) - (4.71) yields as an upper bound for PD_k

$$(4.72) \quad C \log \log n_k (\log n_k)^{-(1+\varepsilon d)},$$

provided k is sufficiently large. Recalling the definition of n_k it is easy to see that (4.72) is summable in k . This completes the proof of (4.67).

A similar but somewhat easier proof can be given for

$$(4.73) \quad \limsup_{n \rightarrow \infty} \sup_{n^{-1} \log n \leq |t| \leq \frac{1}{2}} \frac{|U_{n,v}(t)|}{(\log \log n)^v |t|^{1-v}} = 0 \quad \text{a.s.}$$

Using Fact 1.2 we also have

$$(4.74) \quad \limsup_{n \rightarrow \infty} \sup_{\frac{1}{4} \leq |t| \leq 1} \frac{|U_{n,v}(t)|}{(\log \log n)^v |t|^{1-v}} = 0 \quad \text{a.s.}$$

Combining (4.67), (4.73) and (4.74) completes the upper bound part of the proof.

We now consider the lower bound. It suffices to prove that

$$(4.75) \quad \limsup_{n \rightarrow \infty} \sup_{|t|=a_n} \frac{nF_n(t)}{\log \log n} \geq c\beta_{c/d} \quad \text{a.s.}$$

Let ε be a small positive number and define $\delta = 1 - (1 - \varepsilon)^{1/(2d)}$. We now need the partition $P(\theta)$ defined in (2.22) with $\theta = \theta_n = (1 + o(1))\delta$ such that there exists $\ell = \ell_n \in \mathbb{N}$ with

$$(4.76) \quad \theta^\ell = cn^{-1} \log \log n.$$

This implies $\ell = (1 + o(1))(\log(1/\delta))^{-1} \log n = (1 + o(1))c_1 \log n$, where $c_1 = c_1(\varepsilon) \in (0, \infty)$. Define

$$(4.77) \quad R_n = \{R(a, b) \in P(\theta_n) : |b| = a_n\}.$$

For any $R(a, b) \in R_n$ we have

$$(4.78) \quad |R(a, b)| = (1 - \theta_n)^d a_n = (1 + o(1))(1 - \varepsilon)^{\frac{1}{2}} a_n.$$

Using Fact 4.3 it is easy to see that we have for $\lambda \geq 0$

$$(4.79) \quad P\left(\sup_{|t|=a_n} F_n(t) \geq \lambda\right) \geq P\left(\max_{R(a,b) \in R_n} F_n\{R(a,b)\} \geq \lambda\right)$$

$$\begin{aligned}
&= 1 - P(\max_{R(a,b) \in R_n} F_n\{R(a,b)\} < \lambda) \\
&\geq 1 - \prod_{R(a,b) \in R_n} (1 - P(F_n\{R(a,b)\} \geq \lambda)) \\
&= 1 - (1 - P(F_n\{R\} \geq \lambda))^{\#R_n},
\end{aligned}$$

where R is an arbitrary element of R_n . A combinatorial argument shows

$$(4.80) \quad \#R_n = \binom{\ell+d-1}{d-1} = \binom{(1+o(1))c_1 \log n + d-1}{d-1} = (1+o(1))c_2 (\log n)^{d-1},$$

where $c_2 = c_2(d, \varepsilon) \in (0, \infty)$.

Define $F_{n,m}$ for $m > n$ by

$$(4.81) \quad (m-n)F_{n,m} = mF_m - nF_n$$

and let $\{n_j\}_{j=1}^{\infty}$ be a subsequence of the natural numbers such that $n_j = (1+o(1))(1/(1-(1-\varepsilon)^{\frac{1}{d}}))^j$. Arrived at this stage we can proceed as in Lemma 1, Lemma 2 and the "inner class" proof of Theorem 3 in Kiefer (1972), where (4.75) is proved for $d = 1$. With some minor modifications we obtain for $R \in R_{n_j}$ and $\lambda = (1-\varepsilon)c\beta_{c/d} \log \log n_j$

$$(4.82) \quad P((n_j - n_{j-1})F_{n_{j-1}, n_j}\{R\} \geq \lambda) = j^{-(1+o(1))(1-\varepsilon)d}.$$

Combining this with (4.79) and (4.80) yields

$$\begin{aligned}
(4.83) \quad &P(\sup_{|t|=a_{n_j}} (n_j - n_{j-1})F_{n_{j-1}, n_j}(t) \geq \lambda) \\
&\geq 1 - \{1 - j^{-(1+o(1))(1-\varepsilon)d}\}^{(1+o(1))c_3 j^{d-1}} \\
&= c_3 j^{-(1-d\varepsilon)(1+o(1))},
\end{aligned}$$

where $c_3 = c_3(d, \varepsilon) \in (0, \infty)$. Again like in Kiefer (1972) an application of the Borel-Cantelli lemma completes the proof of (4.75).

□

PROOF OF COROLLARY 4.5. For $0 \leq v < \frac{1}{2}$ this corollary can easily be derived from Theorem 4.3 using Fact 1.2 and (4.75); for $v = \frac{1}{2}$ the result is in Alexander (1984b). It is also easy to derive Corollary 4.5 from Theorem 4.3 for all $0 \leq v \leq \frac{1}{2}$ by use of Corollary 6.2 in Alexander (1982).

□

PROOF OF COROLLARY 4.6. The first part is Theorem 4.3 for $v = 0$, the second part is immediate from the first part and (4.75).

□

3. Suprema over the tail: again laws of the iterated logarithm

Throughout this section let $\{k_n\}_{n=1}^{\infty}$ be any non-decreasing sequence of constants with $0 < k_n \leq n$, $n \in \mathbb{N}$, and define for $d \in \mathbb{N}$ and $0 < v \leq \frac{1}{2}$

$$(4.84) \quad D_{n,v}(k_n) = \sup_{0 < |t| \leq k_n/n} (n/k_n)^v |U_n(t)| / |t|^{\frac{1}{2}-v}$$

and for any fixed positive integer k , $d = 1$ and $0 < v \leq \frac{1}{2}$

$$(4.85) \quad D_{n,v}^{(k)}(k_n) = \begin{cases} \sup_{x_{k:n} < t \leq k_n/n} (n/k_n)^v |U_n(t)| / |t|^{\frac{1}{2}-v} & \text{if } x_{k:n} \leq k_n/n \\ 0 & \text{otherwise;} \end{cases}$$

Note that $1-v$ in the weight functions in sections 1 and 2 is relabeled by $\frac{1}{2}-v$. In the proofs of the theorems in this section we need the following result:

FACT 4.4. (Alexander (1984b); for $d = 1$ see also Kiefer (1972).) If $k_n / \log \log n \rightarrow \infty$ and $\lim_{n \rightarrow \infty} \log \log(n/k_n) / \log \log n = \alpha$, then

$$(4.86) \quad \limsup_{n \rightarrow \infty} \sup_{|t| = k_n/n} n |F_n(t) - |t|| / (k_n \log \log n)^{\frac{1}{2}} \geq (2(1+\alpha(d-1)))^{\frac{1}{2}} \text{ a.s.}$$

For $d = 1$ statement (4.86), with $\alpha(d-1) = 0$, holds true without the second condition.

We now present the theorems of this section for $0 < v < \frac{1}{2}$ followed by their proofs. Then we will consider the case $v = \frac{1}{2}$. For brevity let us write

$$(4.87) \quad a_n = (nk_n^{2v/(1-2v)} (\log \log n)^{1/(1-2v)})^{-1}.$$

THEOREM 4.4. *Let $d \in \mathbb{N}$ and $0 < v < \frac{1}{2}$.*

(I) *If $\sum_{n=1}^{\infty} a_n (\log(1/a_n))^{d-1} = \infty$, then*

$$(4.88) \quad \limsup_{n \rightarrow \infty} D_{n,v}(k_n) / (\log \log n)^{\frac{1}{2}} = \infty \quad \text{a.s.}$$

(II) *If $\sum_{n=1}^{\infty} a_n (\log(1/a_n))^{d-1} < \infty$ and $k_n/n \downarrow 0$, then*

$$(4.89) \quad \limsup_{n \rightarrow \infty} D_{n,v}(k_n) / (\log \log n)^{\frac{1}{2}} \leq (2d)^{\frac{1}{2}} \quad \text{a.s.,}$$

with equality almost surely for the case $d = 1$.

(III) *If in addition to the conditions of part II we assume $\lim_{n \rightarrow \infty} \log \log(n/k_n) / \log \log n = \alpha$, then*

$$(4.90) \quad \limsup_{n \rightarrow \infty} D_{n,v}(k_n) / (\log \log n)^{\frac{1}{2}} = (2(1+\alpha(d-1)))^{\frac{1}{2}} \quad \text{a.s.}$$

THEOREM 4.5. *Let $d = 1$, $k \in \mathbb{N}$ and $0 < v < \frac{1}{2}$.*

(I) *If $\sum_{n=1}^{\infty} n^{k-1} a_n^k = \infty$, then*

$$(4.91) \quad \limsup_{n \rightarrow \infty} D_{n,v}^{(k)}(k_n) / (\log \log n)^{\frac{1}{2}} = \infty \quad \text{a.s.}$$

(II) *If $\sum_{n=1}^{\infty} n^{k-1} a_n^k < \infty$ and $k_n/n \downarrow 0$, then*

$$(4.92) \quad \limsup_{n \rightarrow \infty} D_{n,v}^{(k)}(k_n) / (\log \log n)^{\frac{1}{2}} = 2^{\frac{1}{2}} \quad \text{a.s.}$$

PROOF OF THEOREM 4.4. Part I. By the same reasoning as in the beginning of the proof of Theorem 4.1 we have for any $\varepsilon > 0$: $P(|X|_{1:n} < \varepsilon a_n \text{ i.o.}) = 1$. Now $k_n \uparrow$ implies that $P(|X|_{1:n} < \varepsilon a_n \wedge k_n/n \text{ i.o.}) = 1$. Observe also that $|X|_{1:n} \leq k_n/n$ implies

$$(4.93) \quad \frac{D_{n,v}(k_n)}{(\log \log n)^{\frac{1}{2}}} \geq \left(\frac{n}{k_n}\right)^v \frac{n^{\frac{1}{2}}}{2n(\log \log n)^{\frac{1}{2}} |X|_{1:n}^{\frac{1}{2}-v}}.$$

Hence we have

$$(4.94) \quad \limsup_{n \rightarrow \infty} D_{n,v}(k_n)/(\log \log n)^{\frac{1}{2}} \geq 1/(2\varepsilon^{\frac{1}{2}-v}) \quad \text{a.s.}$$

Letting $\varepsilon \downarrow 0$ proves this part of the theorem.

Part II. Notice that the assumptions of part II imply

$$(4.95) \quad \sum_{n=1}^{\infty} a_n (\log n)^{d-1} < \infty$$

and

$$(4.96) \quad (\log n)^d / (k_n^{2v/(1-2v)} (\log \log n)^{1/(1-2v)}) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Thus for all large n

$$(4.97) \quad D_{n,v}(k_n) \leq E_{n,v}(k_n) + \sup_{b_n \leq |t| \leq k_n/n} (n/k_n)^v |U_n(t)|/|t|^{\frac{1}{2}-v},$$

where

$$(4.98) \quad E_{n,v}(k_n) = \sup_{0 < |t| \leq b_n} (n/k_n)^v |U_n(t)|/|t|^{\frac{1}{2}-v}$$

and $b_n = (\log \log n)^{1/(2v)}/n$.

We now prove

$$(4.99) \quad \lim_{n \rightarrow \infty} E_{n,v}(k_n)/(\log \log n)^{\frac{1}{2}} = 0 \quad \text{a.s.}$$

The proof of (4.99) is similar to the proof of Theorem 4.1, part II, hence we will give a short proof in which for similar quantities as in the proof of Theorem 4.1 the same symbols are used. Define for either choice of sign

$$(4.100) \quad V_n^{\pm} = \sup_{0 < |t| \leq b_n} \frac{\pm(F_n(t) - |t|)}{|t|^{\frac{1}{2}-v}}$$

and let us first show that

$$(4.101) \quad \limsup_{n \rightarrow \infty} \frac{n^{\frac{1}{2}+\nu} V_n^+}{k_n^\nu (\log \log n)^{\frac{1}{2}}} \leq 1 \quad \text{a.s.}, \text{ hence } \lim_{n \rightarrow \infty} \frac{n^{\frac{1}{2}+\nu} V_n^+}{k_n^\nu (\log \log n)^{\frac{1}{2}}} = 0 \quad \text{a.s.}$$

Again we have to prove $\Sigma PC_n < \infty$ and $\lim_{n \rightarrow \infty} PA_n = 0$, where this time

$$(4.102) \quad A_n = \{V_n^+ \geq (\log \log n)^{\frac{1}{2}} k_n^\nu / n^{\frac{1}{2}+\nu}\} \text{ and } C_n = A_n A_{n-1}^c.$$

Let $y_{i,n}$ be the solution of the equation

$$(4.103) \quad f_n(\sigma) = i, \text{ for } i = 0, 1, 2, \dots, i_n = [k_n^\nu (\log \log n)^{1/(2\nu)}],$$

where

$$(4.104) \quad f_n(\sigma) = n\sigma + k_n^\nu (\log \log n)^{\frac{1}{2}} n^{\frac{1}{2}-\nu} \sigma^{\frac{1}{2}-\nu}, \quad \sigma \in [0, \infty),$$

and write again $x_i = x_{i,n} = y_{i,n} \wedge b_n$. For large n we have $C_n \subset \bigcup_{i=1}^{i_n} B_{n,i}$, with $B_{n,i}$ as defined in (4.23), and hence $PC_n \leq \sum_{i=1}^{i_n} PB_{n,i}$. We also have for large n

$$(4.105) \quad (2n)^{-2/(1-2\nu)} \leq x_i \leq i^{2/(1-2\nu)} a_n.$$

Now we can show along the lines of (4.24)-(4.29) that $PC_n = O(a_n (\log n)^{d-1})$ and along the lines of (4.30)-(4.33) that $PA_n = O(n a_n (\log n)^{d-1})$. We omit the straightforward details. Hence (4.101) is established.

Notice that trivially we have

$$(4.106) \quad \frac{n^{\frac{1}{2}+\nu} V_n^-}{k_n^\nu (\log \log n)^{\frac{1}{2}}} \leq \frac{n^{\frac{1}{2}+\nu}}{k_n^\nu (\log \log n)^{\frac{1}{2}}} \frac{(\log \log n)^{(1+2\nu)/(4\nu)}}{n^{\frac{1}{2}+\nu}} \\ = \frac{(\log \log n)^{1/(4\nu)}}{n}.$$

Hence by (4.96)

$$(4.107) \quad \lim_{n \rightarrow \infty} n^{\frac{1}{2}+\nu} V_n^- / (k_n^\nu (\log \log n)^{\frac{1}{2}}) = 0 \quad \text{a.s.},$$

which completes the proof of (4.99).

Now we will consider the second term on the right side of (4.97). This term is for large n less than or equal to

$$(4.108) \quad (\log \log n)^{\frac{1}{2}} (G_n / (\log k_n)^v + H_n),$$

where

$$(4.109) \quad G_n = \sup_{b_n \leq |t| \leq \ell_n/n} |U_n(t)| / (|t| \log \log n)^{\frac{1}{2}},$$

$$(4.110) \quad H_n = \sup_{\ell_n/n \leq |t| \leq k_n/n} |U_n(t)| / (|t| \log \log n)^{\frac{1}{2}},$$

and $\ell_n = k_n / \log k_n$. Notice that $\ell_n/n \downarrow 0$. Using Inequality 2.6 (or Inequality 2.8) in combination with Inequality 2.10 we can prove, similarly to the proof of (4.67) with the v appearing there equal to $\frac{1}{2}$, that

$$(4.111) \quad \limsup_{n \rightarrow \infty} G_n \leq (2(d+1))^{\frac{1}{2}} \text{ a.s. and } \limsup_{n \rightarrow \infty} H_n \leq (2d)^{\frac{1}{2}} \text{ a.s.}$$

From (4.99) and (4.111) we have (4.89). For the case $d = 1$, Fact 4.4 along with (4.89) completes the proof of part II.

Part III. Under the additional assumption of this part we can, again along the lines of the proof of (4.67), improve the result for H_n in (4.111):

$$(4.112) \quad \limsup_{n \rightarrow \infty} H_n \leq (2(1+\alpha(d-1)))^{\frac{1}{2}} \text{ a.s.}$$

Now (4.99), (4.111) and (4.112) yield the upper bound part of (4.90). From Fact 4.4 we can infer the lower bound part. This completes the proof of Theorem 4.4.

□

PROOF OF THEOREM 4.5. Part I. Observe that for any $\varepsilon > 0$,

$\varepsilon a_n \wedge k/(2n) \wedge k_n/n = \varepsilon a_n$ for large n . Combining this with Fact 4.2 we have

$$(4.113) \quad P(X_{k:n} \leq \varepsilon a_n \text{ i.o.}) = 1.$$

As in the proof of Theorem 4.4 this immediately yields (4.91).

Part II. The proof of this part is a mixture of Theorem 4.3, part II, and Theorem 4.2, part II. We now have (4.96) with d replaced by $1/k$ and hence for large n

$$(4.114) \quad D_{n,v}^{(k)}(k_n) \leq E_{n,v}^{(k)}(k_n) + (\log \log n)^{\frac{1}{2}} (G_n / (\log k_n)^v + H_n),$$

where G_n and H_n are defined in (4.109) and (4.110) respectively (of course now $d = 1$),

$$(4.115) \quad E_{n,v}^{(k)}(k_n) = \begin{cases} \sup_{X_{k:n} \leq t \leq b_n} (n/k_n)^v |U_n(t)| / t^{\frac{1}{2}-v} & \text{if } X_{k:n} \leq b_n \\ 0 & \text{otherwise,} \end{cases}$$

and b_n as below (4.98).

Now by the method used in the proof of part II of Theorems 4.1, 4.2 and 4.4 and inclusions like (4.47) - (4.49) it follows that (cf. (4.99))

$$(4.116) \quad \lim_{n \rightarrow \infty} E_{n,v}^{(k)}(k_n) / (\log \log n)^{\frac{1}{2}} = 0 \quad \text{a.s.}$$

But (4.116) and (4.111) with $d = 1$, combined with (4.114), immediately yield

$$(4.117) \quad \limsup_{n \rightarrow \infty} D_{n,v}^{(k)}(k_n) / (\log \log n)^{\frac{1}{2}} \leq 2^{\frac{1}{2}} \quad \text{a.s.}$$

Fact 4.4 along with $P(X_{k:n} > k_n/n \text{ i.o.}) = 0$ completes the proof of this part and hence of Theorem 4.5.

□

Finally we shall deal with the case $v = \frac{1}{2}$, i.e. the unweighted case. But before we present the analogues for $v = \frac{1}{2}$ of Theorems 4.4 and 4.5, we give some results which we shall require in their proofs.

INEQUALITY 4.2. Let $\varepsilon \in (0, \frac{1}{2}]$ and $\alpha \in (0, \frac{1}{4})$. Then we have

$$(4.118) \quad P\left(\sup_{|t| \leq \alpha} |U_n(t)| \geq \lambda \alpha^{\frac{1}{2}}\right) \leq C \left(\log \frac{1}{\alpha}\right)^{d-1} \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2} \psi\left(\frac{\lambda}{(n\alpha)^{\frac{1}{2}}}\right)\right), \quad \lambda \geq 0,$$

where $C = C(d, \varepsilon) \in (0, \infty)$.

The proof of this inequality is similar to, but somewhat easier than, the proof of Inequality 5.1. For this reason we give a full proof of Inequality 5.1 and omit the proof of this inequality.

INEQUALITY 4.3. Let $\varepsilon \in (0,1)$, $\alpha \in (0,1]$ and write $n_k = [(1+\frac{1}{2}\varepsilon)^k]$, $k \in \mathbb{N}$. Then we have for all $k \in \mathbb{N}$ and for $\lambda > 2(\alpha/\varepsilon)^{\frac{1}{2}}$

$$(4.119) \quad P(\max_{n_k < n \leq n_{k+1}} \sup_{|t| \leq \alpha} |U_n(t)| \geq \lambda) \leq 2P(\sup_{|t| \leq \alpha} |U_{n_{k+1}}(t)| \geq (1-\varepsilon)\lambda).$$

PROOF. Similarly as in the proof of Inequality 2.10 we have

$$(4.120) \quad \inf_{n_k < n \leq n_{k+1}; i \in \mathbb{N}} PB_{n,i}^* \cdot P(\sup_{n_k < n \leq n_{k+1}; i \in \mathbb{N}} |S_n(r_i)| > \lambda(n_k+1)^{\frac{1}{2}}) \\ \leq P(\sup_{i \in \mathbb{N}} |S_{n_{k+1}}(r_i)| \geq (1-\frac{1}{2}\varepsilon)\lambda(n_k+1)^{\frac{1}{2}}) \\ \leq P(\sup_{|t| \leq \alpha} |U_{n_{k+1}}(t)| \geq (1-\varepsilon)\lambda),$$

where for $n_k < n \leq n_{k+1}$ and $i \in \mathbb{N}$

$$(4.121) \quad B_{n,i}^* = \{|S_{n_{k+1}}(r_i) - S_n(r_i)| \leq \lambda(n_k+1)^{\frac{1}{2}}\frac{1}{2}\varepsilon\}.$$

So it remains to prove that $PB_{n,i}^* \geq \frac{1}{2}$ for all $n_k < n \leq n_{k+1}$ and all $i \in \mathbb{N}$. But by an application of Chebysev's inequality we have

$$(4.122) \quad PB_{n,i}^{*c} \leq \frac{4|r_i|(1-|r_i|)(n_{k+1}-(n_k+1))}{\lambda^2(n_k+1)\varepsilon^2} \leq \frac{2\varepsilon\alpha}{\lambda^2\varepsilon^2} \leq \frac{1}{2}.$$

□

For the sequence $\{k_n\}_{n=1}^{\infty}$, introduced in the beginning of this section, we have:

FACT 4.5. (Alexander (1984b); for $d = 1$ see also Kiefer (1972).) If $k_n/\log\log n \rightarrow 0$ and $k_n/n \downarrow 0$ then

$$(4.123) \quad \limsup_{n \rightarrow \infty} \sup_{|t|=k_n/n} \frac{n|F_n(t)-|t||}{\log\log n} \log((\log\log n)/k_n) \geq 1 \quad \text{a.s.}$$

THEOREM 4.6. Let $d \in \mathbb{N}$.

(I) If $k_n / \log \log n \rightarrow 0$ and $k_n / n \downarrow 0$, then

$$(4.124) \quad \limsup_{n \rightarrow \infty} D_{n, \frac{1}{2}}(k_n) / (\log \log n)^{\frac{1}{2}} = \infty \quad \text{a.s.}$$

(II) If $k_n / \log \log n \rightarrow c \in (0, \infty)$, then

$$(4.125) \quad \limsup_{n \rightarrow \infty} D_{n, \frac{1}{2}}(k_n) / (\log \log n)^{\frac{1}{2}} = c^{\frac{1}{2}} (\beta_{c/d} - 1) \quad \text{a.s.}$$

(III) If $k_n / \log \log n \rightarrow \infty$ and $k_n / n \downarrow 0$, then

$$(4.126) \quad \limsup_{n \rightarrow \infty} D_{n, \frac{1}{2}}(k_n) / (\log \log n)^{\frac{1}{2}} \leq (2d)^{\frac{1}{2}} \quad \text{a.s.,}$$

with equality almost surely for the case $d = 1$.

(IV) If in addition to the conditions of part III we assume $\lim_{n \rightarrow \infty} \log \log(n/k_n) / \log \log n = \alpha$, then

$$(4.127) \quad \limsup_{n \rightarrow \infty} D_{n, \frac{1}{2}}(k_n) / (\log \log n)^{\frac{1}{2}} = (2(1+\alpha(d-1)))^{\frac{1}{2}} \quad \text{a.s.}$$

THEOREM 4.7. Let $d = 1$ and $k \in \mathbb{N}$.

(I) If $k_n / \log \log n \rightarrow \infty$ and $k_n / n \downarrow 0$, then

$$(4.128) \quad \limsup_{n \rightarrow \infty} D_{n, \frac{1}{2}}^{(k)}(k_n) / (\log \log n)^{\frac{1}{2}} = \infty \quad \text{a.s.}$$

(II) If $k_n / \log \log n \rightarrow c \in (0, \infty)$, then

$$(4.129) \quad \limsup_{n \rightarrow \infty} D_{n, \frac{1}{2}}^{(k)}(k_n) / (\log \log n)^{\frac{1}{2}} = c^{\frac{1}{2}} (\beta_c - 1) \quad \text{a.s.}$$

(III) If $k_n / \log \log n \rightarrow \infty$ and $k_n / n \downarrow 0$, then

$$(4.130) \quad \limsup_{n \rightarrow \infty} D_{n, \frac{1}{2}}^{(k)}(k_n) / (\log \log n)^{\frac{1}{2}} = 2^{\frac{1}{2}} \quad \text{a.s.}$$

PROOF OF THEOREM 4.6. Part I. Assertion (4.124) is an immediate consequence of Fact 4.5.

Part II. First note that we can assume without loss of generality that $k_n = c \log \log n$. Using Inequality 4.2 along with Inequality 4.3 we have, again similarly to the proof of (4.67), that $c^{\frac{1}{2}}(\beta_{c/d} - 1)$ is almost surely an upper bound for the limsup in (4.125). That it is almost surely a lower bound follows from (4.75).

Parts III and IV. Again the upper bounds follow from Inequalities 4.2 and 4.3; the lower bounds follow from Fact 4.4.

□

PROOF OF THEOREM 4.7. Part I. For almost every $\omega \in \Omega$ we can take a subsequence $\{n_j\}_{j=1}^{\infty}$ such that (see Fact 4.5)

$$(4.131) \quad \lim_{j \rightarrow \infty} n_j |F_{n_j}(k_{n_j}/n_j, \omega) - k_{n_j}/n_j| / (k_{n_j} \log \log n_j)^{\frac{1}{2}} = \infty.$$

Now $k_n / \log \log n \rightarrow 0$ implies

$$(4.132) \quad \lim_{j \rightarrow \infty} n_j F_{n_j}(k_{n_j}/n_j, \omega) / (k_{n_j} \log \log n_j)^{\frac{1}{2}} = \infty$$

and hence $X_{k:n_j}(\omega) \leq k_{n_j}/n_j$ for all large j . Thus we have, recalling the definition of $D_{n, \frac{1}{2}}^{(k)}(k_n)$,

$$(4.133) \quad \lim_{j \rightarrow \infty} D_{n_j, \frac{1}{2}}^{(k)}(k_{n_j}, \omega) / (\log \log n_j)^{\frac{1}{2}} = \infty,$$

which completes the proof of this part.

Part II. Again assume $k_n = c \log \log n$. It is well known (e.g. Kiefer (1972)) that for all $\varepsilon \in (0, \infty)$

$$(4.134) \quad P(X_{k:n} \geq ((1+\varepsilon) \log \log n)/n \text{ i.o.}) = 0.$$

From part II of Theorem 4.6 ($d = 1$) and (4.134) combined with (4.75) we obtain part II for $c > 1$. Now we consider $0 < c \leq 1$. From the proof of Theorem 4.6 it follows that for almost every $\omega \in \Omega$ we can take a subsequence $\{n_j\}_{j=1}^{\infty}$ such that

$$(4.135) \quad \lim_{j \rightarrow \infty} n_j |F_{n_j}(k_{n_j}/n_j, \omega) - k_{n_j}/n_j| / (k_{n_j} \log \log n_j)^{\frac{1}{2}} = c^{\frac{1}{2}}(\beta_c - 1).$$

Notice that $0 < c \leq 1$ implies $\beta_c \geq e$. Hence we have from (4.135) that $X_{k:n_j}^{(\omega)} \leq k_{n_j}/n_j$ for all large j , which in combination with (4.135) implies

$$(4.136) \quad \liminf_{j \rightarrow \infty} D_{n_j, \frac{1}{2}}^{(k)}(k_{n_j}, \omega) / (\log \log n_j)^{\frac{1}{2}} \geq c^{\frac{1}{2}} (\beta_c - 1).$$

Now Theorem 4.6, part II, completes the proof of this case and hence of part II.

Part III. This part follows immediately from part II of Theorem 4.6, (4.134) and Fact 4.4.

□

4. Indexation by rectangles

As remarked at the beginning of this chapter, it is not very meaningful to consider the empirical process indexed by rectangles $R \in \bar{\mathcal{R}}$ for the investigations in this chapter. The main reason for this is the fact that the weighted empirical process "blows up" if $|R|$ becomes small, see (3.38). Also, expressions like (4.6) have no meaning if t is replaced by R or, as the case may be, $|R|$. Therefore only the analogues for rectangles of Theorems 4.3 and 4.6 can be considered. However, the analogue of $(k_n/n)^{\frac{1}{2}} D_{n, \frac{1}{2}}(k_n)$, see (4.84), is just the oscillation modulus defined in (5.1) of the next chapter. Hence a theorem for rectangles like Theorem 4.6 will be postponed until Chapter 5, where this oscillation modulus is studied in detail.

We now present analogues of Theorem 4.3 and its corollaries. For this purpose let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive constants such that

$$(4.137) \quad \frac{na_n}{\log n} \rightarrow c \in (0, \infty) \text{ as } n \rightarrow \infty.$$

THEOREM 4.8. *Let $d \in \mathbb{N}$, $0 \leq v \leq \frac{1}{2}$ and $c \in (0, \infty)$. Then we have*

$$(4.138) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |R| \leq 1} \frac{|U_{n,v}\{R\}|}{(\log n)^v |R|^{1-v}} = c^v (\beta_c - 1) \quad \text{a.s.}$$

COROLLARY 4.7. For every fixed $\alpha \in (0,1)$ we have

$$(4.139) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |R| \leq \alpha} \frac{|U_{n,v}\{R\}|}{(\log n)^v (|R|(1-|R|))^{1-v}} = c^v (\beta_c - 1) \quad \text{a.s.}$$

COROLLARY 4.8.

$$(4.140) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |R| \leq 1} |F_n\{R\}/|R|-1| = \beta_c - 1 \quad \text{a.s.,}$$

$$(4.141) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |R| \leq 1} F_n\{R\}/|R| = \beta_c \quad \text{a.s.}$$

PROOF OF THEOREM 4.8. We only give a brief sketch of the proof because it resembles that of Theorem 4.3. Assume without loss of generality that $a_n = cn^{-1} \log n$. By an application of Inequality 2.9, along with Inequality 2.11 and a remark as below (4.68), one can show

$$(4.142) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |R| \leq \frac{1}{2}} \frac{|U_{n,v}\{R\}|}{(\log n)^v |R|^{1-v}} \leq c^v (\beta_c - 1) \quad \text{a.s.}$$

Combining this with Fact 1.2, along with the point representation for rectangles, we see that $c^v (\beta_c - 1)$ is an upper bound for the expression on the left in (4.138) almost surely. In Komlós, Major & Tusnády (1975b, (3.11)) it is shown for $d = 1$, that

$$(4.143) \quad \limsup_{n \rightarrow \infty} \sup_{|R|=a_n} \frac{nF_n\{R\}}{\log n} = c\beta_c \quad \text{a.s.}$$

Now (4.138) readily follows from (4.142) and (4.143). □

PROOF OF COROLLARIES 4.7 AND 4.8. Corollary 4.7 follows easily from Theorem 4.8, Fact 1.2 and (4.143); cf. the proof of Corollary 4.5. The first part of Corollary 4.8 is Theorem 4.8 for $v = 0$, the second part is immediate from the first part and (4.75). □

We conclude this section by presenting a result which is more or less a corollary to our Theorem 4.1 and Corollary 3.9 in Alexander (1984b).

COROLLARY 4.9. Let $d \in \mathbb{N}$ and $\mu \in (-\infty, 1)$. Then we have

(4.144)
$$\limsup_{n \rightarrow \infty} \sup_{|R| \geq \frac{(\log n)^\mu}{n}} \frac{\log \log n}{(\log n)^{1-\frac{1}{2}\mu}} \frac{|U_n\{R\}|}{(|R|(1-|R|))^{\frac{1}{2}}} = \frac{1}{1-\mu} \quad \text{a.s.},$$

whereas

(4.145)
$$\limsup_{n \rightarrow \infty} \sup_{|R| \geq \frac{1}{2}} \frac{1}{(\log n)^{\frac{1}{2}}} \frac{|U_n\{R\}|}{(|R|(1-|R|))^{\frac{1}{2}}} = \infty \quad \text{a.s.}$$

PROOF. "Large d -dimensional rectangles" also behave as "small 1-dimensional points". That means that (4.40) holds true with t replaced by R (cf. the proof of Theorem 3.4). Taking $\nu = \frac{1}{2}$ and combining this with Corollary 3.9 of Alexander (1984b) completes the proof.

□

5. Discussion and bibliography

We begin this section with a diagram which surveys all known strong limit theorems for weighted empirical processes, indexed by points, when the weight functions are of the form $q(\cdot) = \tilde{q}(\cdot) = (\cdot)^\alpha$.

region α	dim.	I^d ; $v = 1-\alpha$		middle of I^d ; $v = 1-\alpha$		tail of I^d ; $v = \frac{1}{2}-\alpha$	
0	$d = 1$ $d \in \mathbb{N}$	Smirnov (1944) Chung (1949) Kiefer (1961)	L	not interesting since LIL holds on whole unit square	L	Einmahl & Mason (1985b)	L
$(0, \frac{1}{2})$	$d = 1$ $d \in \mathbb{N}$	James (1975) Alexander (1982)	L				
$\frac{1}{2}$	$d = 1$ $d \in \mathbb{N}$	Csáki (1974,1975,1982) Einmahl & Mason (1985a)	N	Csáki (1977) Alexander (1984b)	L	immediate from Einmahl & Mason (1985a)	N
$(\frac{1}{2}, 1)$	$d = 1$ $d \in \mathbb{N}$	Mason (1981) Einmahl & Mason (1985a)	N	Einmahl (1984a)	L		
1	$d = 1$ $d \in \mathbb{N}$	Shorack & Wellner (1978) Mason (1982)	N	Wellner (1978) Einmahl (1984a)	L		

L means: LIL (law of the iterated logarithm)
N means: no proper standardization is possible

Theorem 4.1 ($\alpha \in [\frac{1}{2}, 1]$), which is contained in Einmahl & Mason (1985a), is a generalization of various results in the literature in which special values of v and d are considered. Corollary 4.1 shows that for $V_{n,v}$ no LIL type result holds. However, by Corollary 4.2, $\log V_{n,v}$ obeys a LIL. Note that Corollary 4.3 implies a Glivenko-Cantelli theorem for the weighted empirical process if $0 < v \leq \frac{1}{2}$ (and, of course, for $\frac{1}{2} < v \leq 1$). Theorem 4.2 is contained in Einmahl, Häusler and Mason (1985). It shows another way of generalizing Theorem 4.1 with $d = 1$. We remark that since it can be inferred from Fact 4.1 that for any sequence of positive constants $\{b_n\}_{n=1}^{\infty}$, we have

$$\sum_{n=1}^{\infty} b_n = \infty \text{ implies } P(X_{1:n} \leq b_n \wedge 1/(2n) \text{ i.o.}) = 1,$$

it easily follows that part I of this theorem, for $k = 1$, holds true without the condition $a_n \downarrow$.

Theorem 4.3 ($\alpha \in [\frac{1}{2}, 1]$) is contained in Einmahl (1984a) and is also considered by several authors for special values of v and d . This theorem shows that a LIL does hold if we don't take the suprema over the whole unit square; cf. Theorem 4.1. A closer inspection of the number on the right in (4.63) even shows that $\{\text{cloglog } n/n\}_{n=1}^{\infty}$ is, roughly speaking, the smallest sequence for which a LIL, in this sense, can be proved; see Alexander (1984b, p.3) for properties of β_{σ} . As far as this theorem is concerned no results were known for v strictly between 0 and $\frac{1}{2}$, not even for $d = 1$. Alexander (1984d) pointed out to me, however, that the result for $0 \leq v < \frac{1}{2}$ in Theorem 4.3 can be obtained from the result for $v = \frac{1}{2}$ and its proof, instead of giving a direct proof. This remark can also be made for Theorem 4.8. It should be noted that the sequence $\{\text{cloglog } n/n\}_{n=1}^{\infty}$ plays a similar role for the empirical process, indexed by points, as $\{\text{clog } n/n\}_{n=1}^{\infty}$ does for the empirical process, indexed by rectangles; see e.g. Theorems 4.3, 4.6, 4.8 and 5.3. It is also interesting to observe that there is a discontinuity in behaviour in $v = \frac{1}{2}$ for the limsup in (4.63) if $c^{\frac{1}{2}}(\beta_{c/d} - 1) < (2(d+1))^{\frac{1}{2}}$. The results in Corollary 4.6 are the multivariate versions of Corollary 2(iii) and Theorem 2(i), respectively, in Wellner (1978).

For values of α in $[0, \frac{1}{2})$ the results differ from the results in sections 1 and 2. In that case it is shown in James (1975, $d = 1$) and in Alexander (1982, Corollary 6.2, $d \in \mathbb{N}$) that a LIL holds already, when the suprema are taken over the whole unit square I^d . Therefore it is not

interesting to consider an analogue of Theorem 4.3 for these values of α .

Theorems 4.4-4.7 ($\alpha \in [0, \frac{1}{2})$) are established in Einmahl & Mason (1985b). As far as I know, this is the first time that results of this type have been considered. These theorems show that for the given weight functions a LIL in the tail holds under additional assumptions on the sequence $\{k_n\}_{n=1}^{\infty}$. This is not true if $D_{n,v}(k_n)$ and $D_{n,v}^{(k)}(k_n)$ are considered for $v \in [-\frac{1}{2}, 0]$, because it readily follows from Theorems 4.1 and 4.2 and their proofs that in those cases we have again the " ∞ or 0 behaviour". If we consider Theorem 4.4, for convenience we choose $d = 1$, or Theorem 4.5 we see that the value of the limsup is "always" ∞ or $2^{\frac{1}{2}}$. The proofs of Theorems 4.4 and 4.5 show indeed what this remark suggests, namely that these theorems describe a behaviour which is a kind of mixture of that of Theorems 4.1-4.2 and Theorem 4.3. Note that in contrast to Theorems 4.4 and 4.5, the values of the limsup in Theorems 4.6 and 4.7 decrease continuously from ∞ to $2^{\frac{1}{2}}$ as the order of magnitude of the sequences $\{k_n\}_{n=1}^{\infty}$ increases.

Theorem 4.8, for $v = \frac{1}{2}$ and $d \in \mathbb{N}$, is obtained in Alexander (1984b, Corollary 3.9). In that corollary also truncation points other than $\log n/n$ are considered; see also Mason, Shorack & Wellner (1983) for $d = 1$. Corollary 4.9, contained in Einmahl & Mason (1985a), nicely illustrates that in one situation small rectangles determine the value of the limsup (or lim), whereas in a slightly different situation, large rectangles do. For distributional analogues of the results in this chapter when $d = 1$, the reader is referred to Jaeschke (1979), Eicker (1979), Mason (1983, 1985) and Csörgö & Mason (1985).

CHAPTER 5

LIMIT THEOREMS FOR THE OSCILLATION MODULI
OF THE MULTIVARIATE UNIFORM EMPIRICAL PROCESS

First we give the definitions of the two different oscillation moduli of the multivariate empirical process which we consider; in the one-dimensional case these definitions (almost) coincide. The oscillation modulus with which we shall deal mostly is

$$(5.1) \quad \omega_n(\alpha) = \sup_{R \in \mathcal{R}; |R| \leq \alpha} |U_n\{R\}|, \quad 0 \leq \alpha \leq 1;$$

slight modifications of ω_n have applications in the theory of density estimation. We will also consider the oscillation modulus

$$(5.2) \quad M_n(\alpha) = \sup_{|t_j - s_j| \leq \alpha, \forall j} |U_n(t) - U_n(s)|, \quad \alpha \geq 0;$$

this modulus can be used for proving tightness of the sequence $\{U_n\}_{n=1}^{\infty}$. In section 1 we derive the exact order of magnitude of the moments of both $\omega_n(\alpha)$ and $M_n(\alpha)$ for n somewhat larger than $1/\alpha$. We consider, in section 2, the almost sure behaviour of $\omega_n(a_n)$, where $\{a_n\}_{n=1}^{\infty}$ are sequences of positive numbers decreasing to 0. These sequences may tend to 0 at all possible rates, i.e. we consider sequences which are almost constant, sequences which are so small that almost surely $\sup_{|R| \leq a_n} nF_n\{R\}$ tends to 1 and everything in between.

1. The order of magnitude of the moments

We begin this section by giving a sharp upper bound and a lower bound for the probability that $\omega_n(\alpha)$ exceeds a certain value. Using these bounds the order of magnitude of the moments of $\omega_n(\alpha)$ will follow easily from the relation

$$(5.3) \quad EX^v = \int_0^{\infty} P(X > \lambda^{1/v}) d\lambda, \quad v \in \mathbb{N},$$

where X is a non-negative rv. In the second part of this section the same will be done for M_n .

INEQUALITY 5.1. Let $\varepsilon \in (0, \frac{1}{2}]$ and $\alpha \in (0, \frac{1}{2})$. Then we have

$$(5.4) \quad P(\omega_n(\alpha) \geq \lambda \alpha^{\frac{1}{2}}) \leq C \frac{1}{\alpha} (\log \frac{1}{\alpha})^{d-1} \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2}\right) \psi\left(\frac{\lambda}{(n\alpha)^{\frac{1}{2}}}\right), \quad \lambda \geq 0,$$

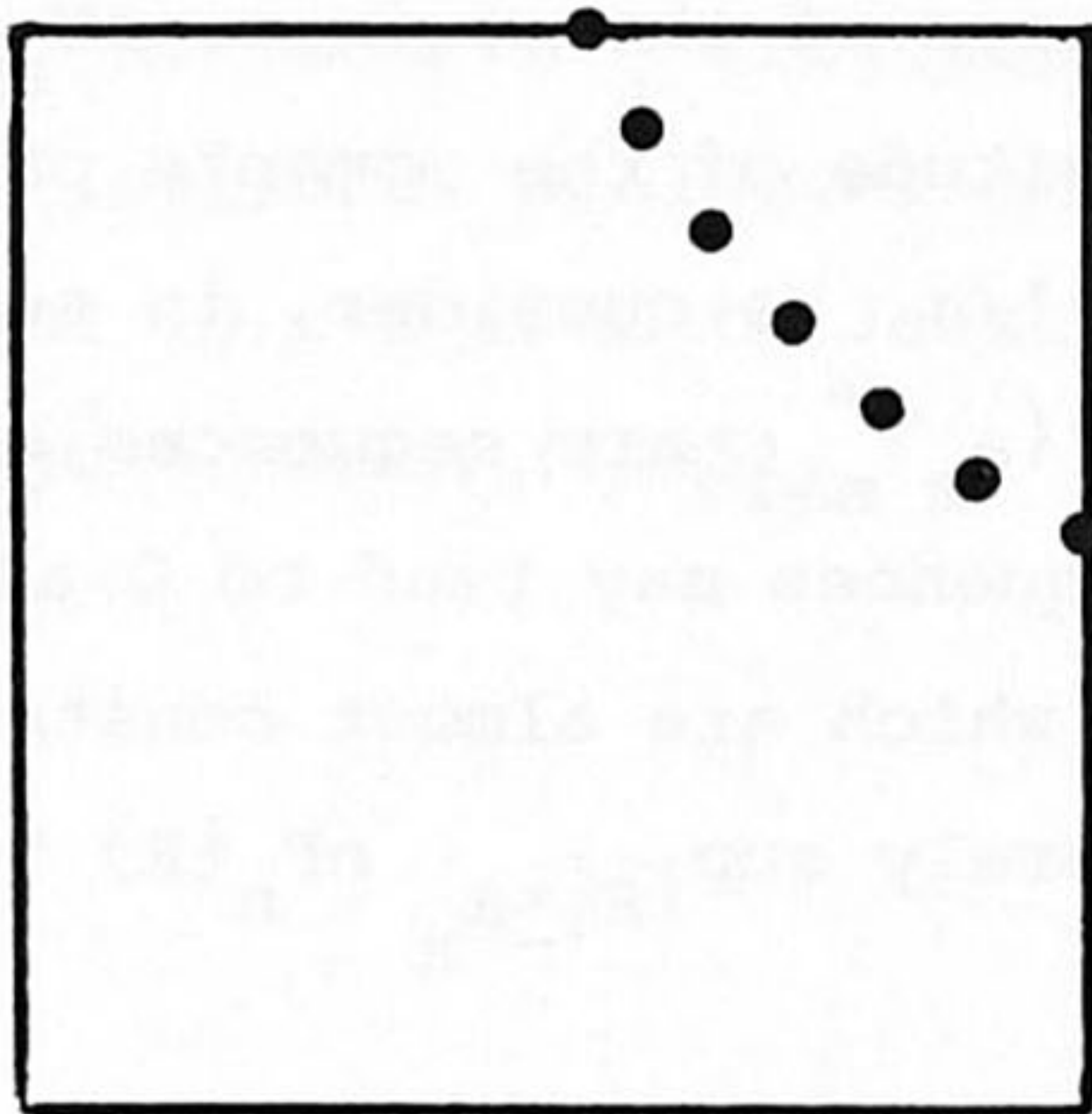
where $C = C(d, \varepsilon) \in (0, \infty)$.

PROOF. For this proof we use the point representation for rectangles, as defined in (2.34); we then have with the notation used there

$$(5.5) \quad P(\omega_n(\alpha) \geq \lambda \alpha^{\frac{1}{2}}) = P\left(\sup_{\bar{F}(\langle s, t \rangle) \leq \alpha} |\bar{U}_n(\langle s, t \rangle)| \geq \lambda \alpha^{\frac{1}{2}}\right).$$

Next we need some definitions. Define θ by $\theta^{2d+1} = 1 - \varepsilon$ and $\ell \in \mathbb{N}$ by $\theta^{\ell+1} < \alpha \leq \theta^\ell$ and note that $\ell \leq \log \frac{1}{\alpha} / \log \frac{1}{\theta}$. For $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ define

$$(5.6) \quad A(\theta, k) = \{u \in I^d : \langle u_j \rangle = \langle \theta^{k(j)} \rangle, (k(j) \in \mathbb{N}_0) \text{ and } \sum_{j=1}^d k(j) = k\}$$

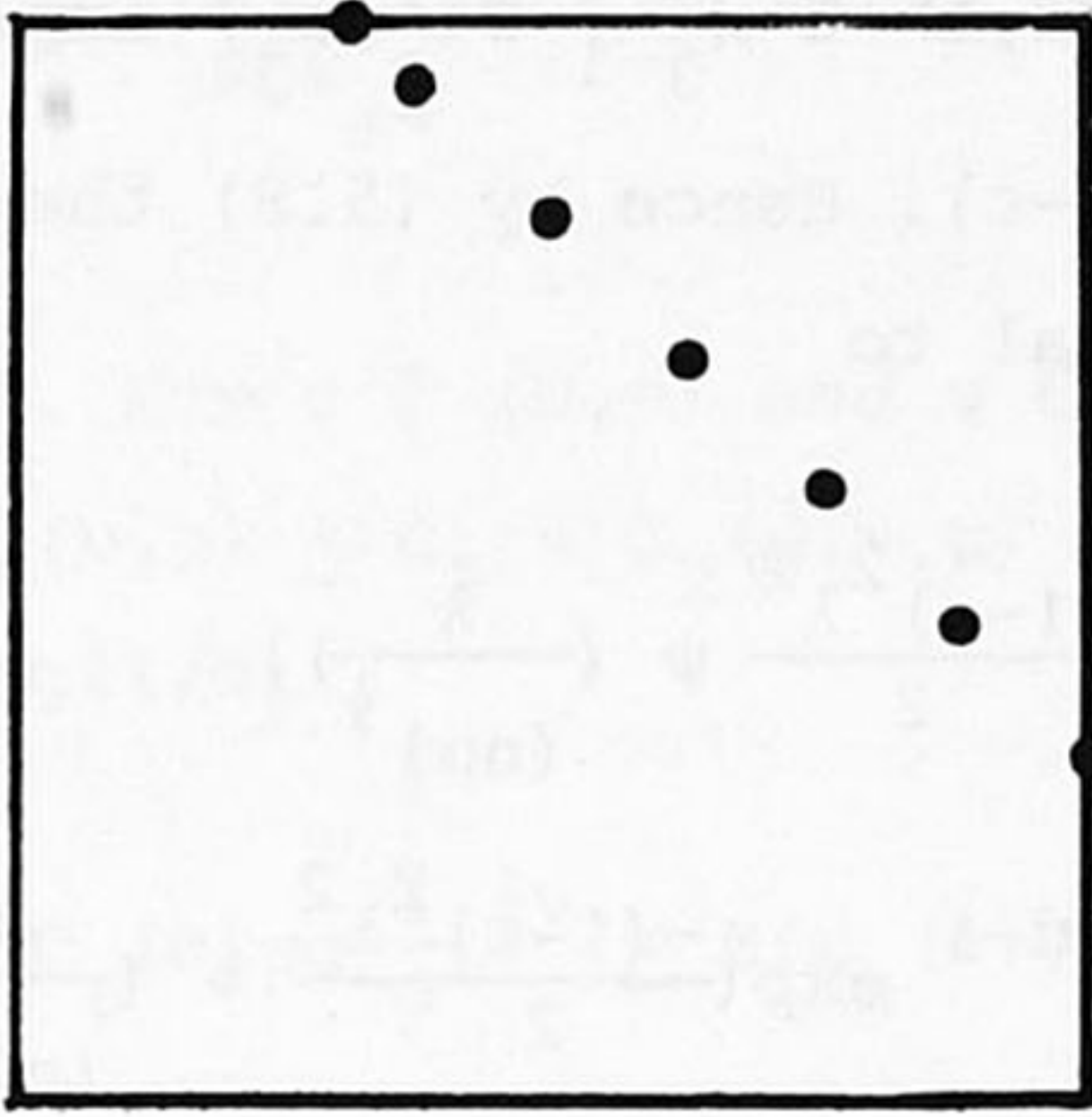


$$d = 2$$

$$A\left(\frac{9}{10}, 6\right)$$

and for $\beta \in (1, 2)$ and $x \in (0, 1)$

$$(5.7) \quad B(\beta, x) = \begin{cases} \{\langle 1 - k(\beta - 1)x, \beta x + k(\beta - 1)x \rangle \in I^2 : k \in \mathbb{N}_0, k \leq \left[\frac{1 - \beta x}{(\beta - 1)x}\right]\} \\ \cup \{\langle \beta x, 1 \rangle\} & \text{if } \beta x < 1 \\ \{\langle 1, 1 \rangle\} & \text{if } \beta x \geq 1. \end{cases}$$



$$B(1\frac{2}{3}, \frac{3}{16})$$

It is easily seen that $\#A(\theta, k) = \binom{k+d-1}{d-1}$, cf. (4.80), and that $\#B(\beta, x) \leq 1/((\beta-1)x)$.

From (2.19) and the fourth expression in (2.15) it follows that for $\langle \tilde{s}, \tilde{t} \rangle \in I^{2d}$ ($\tilde{s}, \tilde{t} \in I^d$) with $\bar{F}(\langle \tilde{s}, \tilde{t} \rangle) \leq \alpha/(1-\epsilon)$

$$(5.8) \quad P(\sup_{\langle s, t \rangle \leq \langle \tilde{s}, \tilde{t} \rangle} |\bar{U}_n(\langle s, t \rangle)| \geq \lambda \alpha^{\frac{1}{2}}) \\ \leq C \exp\left(\frac{-(1-\epsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(n\alpha)^{\frac{1}{2}}}\right)\right).$$

Using (5.5) - (5.7) we have

$$(5.9) \quad P(\omega_n(\alpha) \geq \lambda \alpha^{\frac{1}{2}}) \leq P(\sup_{\bar{F}(\langle s, t \rangle) \leq \theta^\ell} |\bar{U}_n(\langle s, t \rangle)| \geq \lambda \alpha^{\frac{1}{2}}) \\ \leq P(\max_{u \in A(\theta, \ell-d)} \sup_{\langle t_j + s_j - 1 \rangle \leq u} |\bar{U}_n(\langle s, t \rangle)| \geq \lambda \alpha^{\frac{1}{2}}) \\ \leq \sum_{u \in A(\theta, \ell-d)} P(\sup_{\langle t_j + s_j - 1 \rangle \leq u} |\bar{U}_n(\langle s, t \rangle)| \geq \lambda \alpha^{\frac{1}{2}}) \\ \leq \sum_{u \in A(\theta, \ell-d)} P(\max_{\substack{\langle \tilde{s}, \tilde{t} \rangle \in \Pi \\ j=1}}^d \sup_{B(\frac{1}{\theta}, u_j)} |\bar{U}_n(\langle s, t \rangle)| \geq \lambda \alpha^{\frac{1}{2}}) \\ \leq \sum_{u \in A(\theta, \ell-d)} \sum_{\substack{\langle \tilde{s}, \tilde{t} \rangle \in \Pi \\ j=1}}^d P(\sup_{B(\frac{1}{\theta}, u_j)} |\bar{U}_n(\langle s, t \rangle)| \geq \lambda \alpha^{\frac{1}{2}}),$$

where $\prod_{j=1}^d B(\frac{1}{\theta}, u_j)$ is the Cartesian product of the $B(\frac{1}{\theta}, u_j)$, formed in the same way as below (2.39). Notice that $\langle \tilde{s}, \tilde{t} \rangle \in \prod_{j=1}^d B(\frac{1}{\theta}, u_j)$ implies $\bar{F}(\langle \tilde{s}, \tilde{t} \rangle) = \prod_{j=1}^d (\frac{u_j}{\theta} \wedge 1) \leq \theta^{\ell-2d} \leq \alpha/(1-\epsilon)$. Hence by (5.8) the last expression in (5.9) is less than or equal to

$$\begin{aligned} & \binom{\ell-1}{d-1} \prod_{j=1}^d \left(\frac{\theta}{(1-\theta)u_j} \right) C \exp\left(\frac{-(1-\epsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(n\alpha)^{\frac{1}{2}}} \right) \right) \\ & \leq C (\log \frac{1}{\theta})^{1-d} \left(\frac{\theta}{1-\theta} \right)^d \frac{1}{\alpha} (\log \frac{1}{\alpha})^{d-1} \exp\left(\frac{-(1-\epsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(n\alpha)^{\frac{1}{2}}} \right) \right). \end{aligned}$$

Relabeling $(1-\epsilon)^2$ by $(1-\epsilon)$ (which entails $\theta^{4d+2} = 1-\epsilon$) and $C(\log(1/\theta))^{1-d} (\theta/(1-\theta))^d$ by C , we obtain the right side of (5.4). □

INEQUALITY 5.2. Let $\alpha \in (0, \frac{1}{2}]$ and $\mu \in (0, \infty)$. Then we have for $n \geq (\mu/\alpha) \log(1/\alpha)$ and $c_1 \alpha^{\frac{1}{2}} \leq \lambda \leq c_2 \alpha^{\frac{1}{2}} (\log(1/\alpha))^{\frac{1}{2}}$

$$(5.10) \quad P(\omega_n(\alpha) > \lambda) \geq 1 - (1 - \exp(-c_3 \lambda^2/\alpha))^{[1/\alpha]},$$

where $c_2 = c_2(\mu)$ and $c_1, c_2, c_3 \in (0, \infty)$.

For the proof of Inequality 5.2 we need

FACT 5.1. (Kolmogorov (1929); see also Alexander (1984b), Lemma 7.5.) Let $R \in \bar{R}$ with $|R| \leq \frac{1}{2}$. Then for every $\epsilon > 0$ there exist $c_4, c_5 \in (0, \infty)$ such that for $c_4 |R|^{\frac{1}{2}} \leq \lambda \leq c_5 n^{\frac{1}{2}} |R|$

$$(5.11) \quad P(U_n\{R\} > \lambda) \geq \exp(-(1+\epsilon)\lambda^2/(2|R|(1-|R|))).$$

PROOF OF INEQUALITY 5.2. Let \bar{R}_α be a subset of \bar{R} consisting of $[1/\alpha]$ rectangles R with disjoint interiors and $|R| = \alpha$. Using Fact 4.3 and Fact 5.1, with e.g. $\epsilon = 1$, now immediately yields

$$\begin{aligned} (5.12) \quad P(\omega_n(\alpha) > \lambda) & \geq P(\max_{R \in \bar{R}_\alpha} U_n\{R\} > \lambda) \\ & = 1 - P(\max_{R \in \bar{R}_\alpha} U_n\{R\} \leq \lambda) \geq 1 - \prod_{R \in \bar{R}_\alpha} P(U_n\{R\} \leq \lambda) \end{aligned}$$

$$= 1 - \prod_{R \in \bar{R}_\alpha} (1 - P(U_n\{R\} > \lambda)) \geq 1 - (1 - \exp(-c_3 \frac{\lambda^2}{\alpha}))^{[1/\alpha]}.$$

□

THEOREM 5.1. For $\mu \in (0, \infty)$ and $v \in \mathbb{N}$ there exist numbers

$0 < C_1 = C_1(v, \mu) \leq C_2 = C_2(d, v, \mu) < \infty$ such that for all $\alpha \in (0, \frac{1}{2}]$ and $n \geq (\mu/\alpha) \log(1/\alpha)$

$$(5.13) \quad C_1 (\alpha \log \frac{1}{\alpha})^{\frac{1}{2}v} \leq E(\omega_n(\alpha))^v \leq C_2 (\alpha \log \frac{1}{\alpha})^{\frac{1}{2}v}.$$

PROOF. First observe that we may (and will) assume without loss of generality that $1/\alpha \in \mathbb{N}$, α is sufficiently small and $\mu \leq 1$. Write $m = 1/\alpha$. Then we have for the upper bound by (5.3) and Inequality 5.1 with $\varepsilon = \frac{1}{2}$

$$(5.14) \quad E(\omega_n(\alpha))^v \leq c_4 \left(\frac{\log m}{m}\right)^{\frac{1}{2}v} \leq \int_0^\infty 1 d\lambda + \int_{c_4 \left(\frac{\log m}{m}\right)^{\frac{1}{2}v}}^\infty C m (\log m)^{d-1} \exp\left(-\frac{\lambda^{2/v} m}{4} \psi\left(\frac{\lambda^{1/v} m}{n^{\frac{1}{2}}}\right)\right) d\lambda$$

$$\leq c_4 \left(\frac{\log m}{m}\right)^{\frac{1}{2}v} + C m (\log m)^{d-1} \int_{c_4 \left(\frac{\log m}{m}\right)^{\frac{1}{2}v}}^\infty \exp\left(-\frac{\lambda^{2/v} m}{4} \psi\left(\frac{\lambda^{1/v} m^{\frac{1}{2}}}{\mu^{\frac{1}{2}} (\log m)^{\frac{1}{2}}}\right)\right) d\lambda,$$

where $c_4 = 8^v \mu^{-\frac{1}{2}v}$. We now see by the change of variables $m^{\frac{1}{2}} \lambda^{1/v} = \mu^{\frac{1}{2}} (\log m)^{\frac{1}{2}} y$ that this last integral is equal to

$$(5.15) \quad \nu \mu^{\frac{1}{2}v} \left(\frac{\log m}{m}\right)^{\frac{1}{2}v} \int_{8/\mu}^\infty \exp(-y^2 \log m \frac{\mu}{4} \psi(y)) y^{v-1} dy.$$

From (2.15) it follows that $y^2 \psi(y) \geq y$ for $y \geq 8$. Hence we have by elementary analysis that the expression in (5.15) is less than or equal to $c_5(v, \mu) \left(\frac{\log m}{m}\right)^{\frac{1}{2}v} \frac{1}{m^2 \log m}$. This in combination with inequality (5.14) completes the proof of the upper bound part.

Now we consider the lower bound. By Inequality 5.2 and again (5.3) we have

$$(5.16) \quad E(\omega_n(\alpha))^v \geq \int_{c_1 m^{-\frac{1}{2}v}}^{c_2 \left(\frac{\log m}{m}\right)^{\frac{1}{2}v}} 1 - (1 - \exp(-c_3 m \lambda^{2/v}))^m d\lambda$$

$$\geq \int_{\frac{c_1^v}{c_2^v \wedge c_3^{-\frac{1}{2}v}}}^{\frac{c_2^v \wedge c_3^{-\frac{1}{2}v}}{(\frac{\log m}{m})^{\frac{1}{2}v}}} 1 - (1 - \frac{1}{m})^m d\lambda$$

$$\geq \frac{1}{2} (c_2^v \wedge c_3^{-\frac{1}{2}v}) (\frac{\log m}{m})^{\frac{1}{2}v} (1 - \frac{1}{e}),$$

where this last inequality holds for large m . This completes the proof of the theorem. □

In the remainder of this section we will derive the order of magnitude of the moments of $M_n(\alpha)$. Actually we will prove exactly the same theorem as we proved for $\omega_n(\alpha)$. To achieve this we need similar inequalities for M_n as the Inequalities 5.1 and 5.2. To facilitate their proofs we will present these inequalities under the inessential condition that $1/\alpha \in \mathbb{N}$.

We now give some notation which is needed in these proofs. For $1/\alpha \in \mathbb{N}$, $P_\alpha \subset R$ is the partition of $(0,1]^d$ into the squares of equal size

$$(5.18) \quad R_{k(1), \dots, k(d)} = ((k(1)-1)\alpha, k(1)\alpha] \times \dots \times ((k(d)-1)\alpha, k(d)\alpha],$$

with $k(j) \in \mathbb{N}$ and $k(j) \leq 1/\alpha$ for all $j \leq d$. Let us also introduce for $j \leq d$ and $1 \leq k \leq 1/\alpha$ ($k \in \mathbb{N}$) the slices

$$(5.19) \quad C_{j,k} = (0,1]^{j-1} \times ((k-1)\alpha, k\alpha] \times (0,1]^{d-j}$$

and the rv's

$$(5.20) \quad v_{j,k} = \sup_{S \subset C_{j,k}; S \in R} |U_n\{S\}|.$$

Finally we define an auxiliary oscillation modulus by

$$(5.21) \quad \tilde{M}_n(\alpha) = \max_{R \in P_\alpha} \sup_{s, t \in R} |U_n(t) - U_n(s)|, \quad 1/\alpha \in \mathbb{N};$$

observe that since $U_n \in D_d$ we have

$$(5.22) \quad \tilde{M}_n(\alpha) \leq M_n(\alpha) \leq 2^d \tilde{M}_n(\alpha), \quad 1/\alpha \in \mathbb{N}.$$

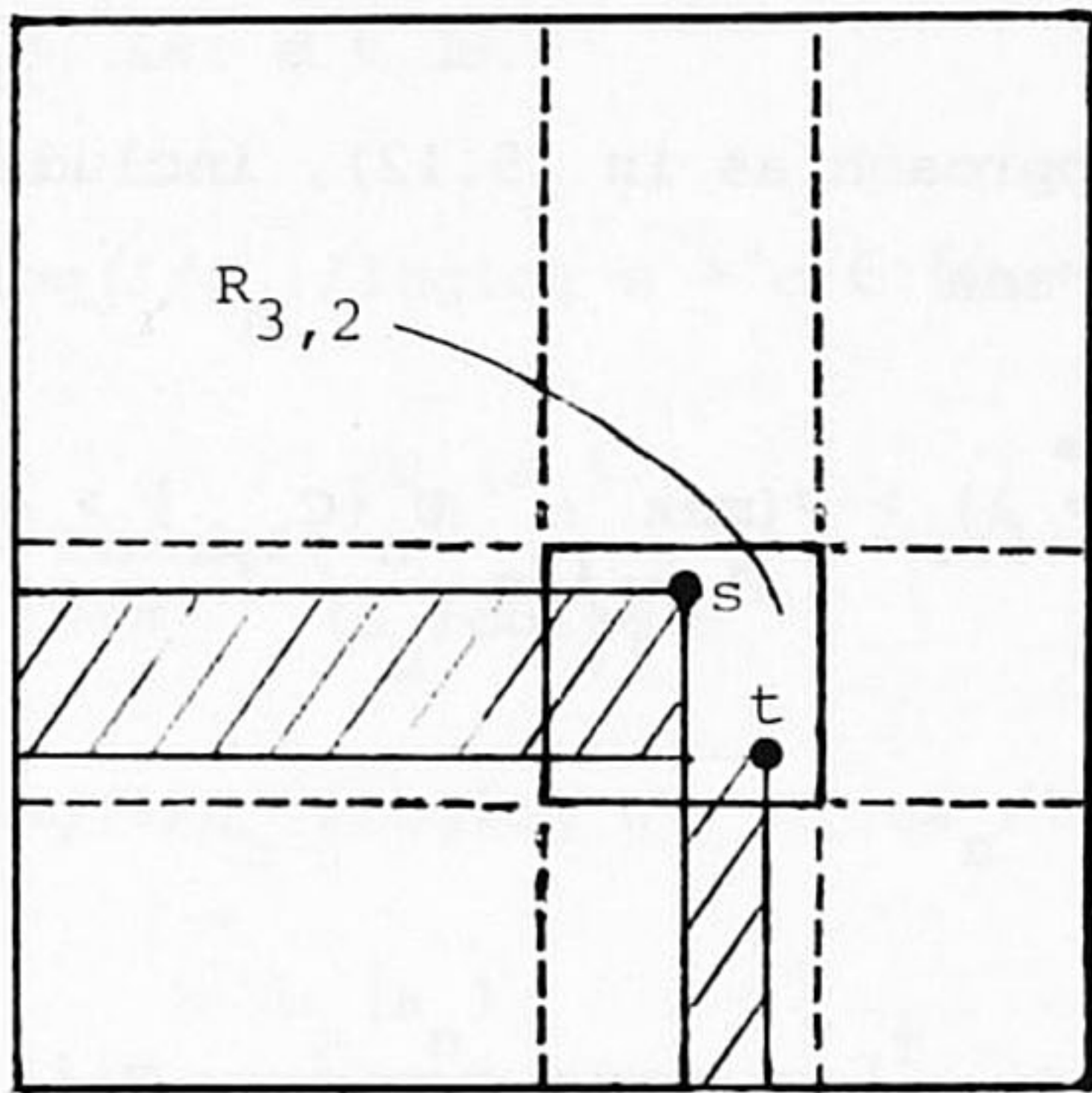
INEQUALITY 5.3. For $\alpha \in (0, \frac{1}{2}]$ and $1/\alpha \in \mathbb{N}$ we have

$$(5.23) \quad P(M_n(\alpha) \geq \lambda) \leq C \frac{1}{\alpha} \exp(-c_1 \frac{\lambda^2}{\alpha} \psi(\frac{\lambda}{n^{\frac{1}{2}} \alpha})), \quad \lambda \geq 0,$$

where $C = C(d)$, $c_1 = c_1(d)$ and $C, c_1 \in (0, \infty)$.

PROOF. First note that

$$(5.24) \quad \sup_{s, t \in R_{k(1), \dots, k(d)}} |U_n(t) - U_n(s)| \leq \sum_{j=1}^d V_{j, k(j)}.$$



$$d = 2$$

$$P(\frac{1}{4})$$

Hence we have

$$(5.25) \quad \begin{aligned} \tilde{M}_n(\alpha) &\leq \max_{k(1), \dots, k(d)} \sum_{j=1}^d V_{j, k(j)} \\ &\leq \sum_{j=1}^d \max_{k(1), \dots, k(d)} V_{j, k(j)} = \sum_{j=1}^d \max_{1 \leq k \leq 1/\alpha} V_{j, k}. \end{aligned}$$

By combining (5.22) and (5.25) and applying (2.19) with $\epsilon = \frac{1}{2}$ we obtain

$$(5.26) \quad \begin{aligned} P(M_n(\alpha) \geq \lambda) &\leq P(\sum_{j=1}^d \max_{1 \leq k \leq 1/\alpha} V_{j, k} \geq \lambda/2^d) \\ &\leq \sum_{j=1}^d P(\max_{1 \leq k \leq 1/\alpha} V_{j, k} \geq \lambda/(d2^d)) \end{aligned}$$

$$\leq \sum_{j=1}^d \sum_{k=1}^{1/\alpha} P(V_{j,k} \geq \lambda / (d2^d)) \leq dC \frac{1}{\alpha} \exp(-c_1 \frac{\lambda^2}{\alpha} \psi(\frac{\lambda}{n^{\frac{1}{2}\alpha}})),$$

where for the last inequality the fact that ψ decreases is used. Relabeling dC by C completes the proof. □

INEQUALITY 5.4. Let $\alpha \in (0, \frac{1}{2}]$, $1/\alpha \in \mathbb{N}$ and $\mu \in (0, \infty)$. Then we have for $n \geq (\mu/\alpha) \log(1/\alpha)$ and $c_1 \alpha^{\frac{1}{2}} \leq \lambda \leq c_2 \alpha^{\frac{1}{2}} (\log(1/\alpha))^{\frac{1}{2}}$

$$(5.27) \quad P(M_n(\alpha) > \lambda) \geq 1 - (1 - \exp(-c_3 \lambda^2 / \alpha))^{1/\alpha},$$

where $c_2 = c_2(\mu)$ and $c_1, c_2, c_3 \in (0, \infty)$.

PROOF. Using (5.22) and the same approach as in (5.12), including Facts 4.3 and 5.1, we immediately see

$$(5.28) \quad P(M_n(\alpha) > \lambda) \geq P(\tilde{M}_n(\alpha) > \lambda) \geq P(\max_{1 \leq k \leq 1/\alpha} U_n\{C_{1,k}\} > \lambda) \\ \geq 1 - (1 - \exp(-c_3 \frac{\lambda^2}{\alpha}))^{1/\alpha}.$$

□

We now present our moment theorem for $M_n(\alpha)$, the proof of which will be omitted because it is very much the same as the proof of Theorem 5.1. This can easily be seen by comparing the Inequalities 5.1 and 5.2 and the Inequalities 5.3 and 5.4.

THEOREM 5.2. For $\mu \in (0, \infty)$ and $v \in \mathbb{N}$ there exist numbers

$0 < C_1 = C_1(v, \mu) \leq C_2 = C_2(d, v, \mu) < \infty$ such that for all $\alpha \in (0, \frac{1}{2}]$ and $n \geq (\mu/\alpha) \log(1/\alpha)$

$$(5.29) \quad C_1 (\alpha \log \frac{1}{\alpha})^{\frac{1}{2}v} \leq E(M_n(\alpha))^v \leq C_2 (\alpha \log \frac{1}{\alpha})^{\frac{1}{2}v}.$$

2. The almost sure behaviour of the oscillation modulus ω_n

Just as in Chapter 4 we need for the strong results in this section a maximal inequality. Using again the point representation for rectangles its

proof is, mutatis mutandis, identical to the proof of Inequality 4.3 (see also the remark below Inequality 2.10).

INEQUALITY 5.5. Let $\varepsilon \in (0,1)$, $\alpha \in (0,1]$ and write $n_k = [(1+\frac{1}{2}\varepsilon)^k]$, $k \in \mathbb{N}$. Then we have for all $k \in \mathbb{N}$ and for $\lambda > 2(\alpha/\varepsilon)^{\frac{1}{2}}$

$$(5.30) \quad P(\max_{n_k < n \leq n_{k+1}} \omega_n(\alpha) \geq \lambda) \leq 2P(\omega_{n_{k+1}}(\alpha) \geq (1-\varepsilon)\lambda).$$

Now we are ready to present and prove the theorem of this section. For this purpose let $\{a_n\}_{n=1}^{\infty}$ be a sequence of numbers in $(0,1]$ decreasing to 0 and let us recall the definition of β_σ in (4.62).

THEOREM 5.3. Let $d \in \mathbb{N}$.

(I) If $\log(1/a_n)/\log \log n \rightarrow c \in [0, \infty)$ and $na_n \uparrow$ if $c = 0$, then

$$(5.31) \quad \limsup_{n \rightarrow \infty} \frac{\omega_n(a_n)}{(a_n \log \log n)^{\frac{1}{2}}} = (2(1+c))^{\frac{1}{2}} \quad \text{a.s.}$$

(II) If $\log(1/a_n)/\log \log n \rightarrow \infty$, $na_n/\log n \rightarrow \infty$ and $na_n \uparrow$, then

$$(5.32) \quad \lim_{n \rightarrow \infty} \frac{\omega_n(a_n)}{(a_n \log(1/a_n))^{\frac{1}{2}}} = 2^{\frac{1}{2}} \quad \text{a.s.}$$

(III) If $na_n/\log n \rightarrow c \in (0, \infty)$, then

$$(5.33) \quad \lim_{n \rightarrow \infty} \frac{n^{\frac{1}{2}} \omega_n(a_n)}{\log n} = c(\beta_c - 1) \quad \text{a.s.}$$

(IV) If $na_n/\log n \rightarrow 0$ and $\log(1/a_n)/\log n \rightarrow 1$, then

$$(5.34) \quad \lim_{n \rightarrow \infty} \frac{n^{\frac{1}{2}} \omega_n(a_n) \log((\log n)/(na_n))}{\log n} = 1 \quad \text{a.s.}$$

(V) If $\log(1/a_n)/\log n = c \in (1, \infty)$ (i.e. $a_n = n^{-c}$), then

$$(5.35) \quad \limsup_{n \rightarrow \infty} n^{\frac{1}{2}} \omega_n(a_n) = [c/(c-1)] \quad \text{a.s.,}$$

if in addition $c/(c-1) \notin \mathbb{N}$, then c may be replaced by $c_n \rightarrow c$ and \limsup by \lim .

PROOF. Let us first prove that the numbers on the right are upper bounds for the respective expressions on the left almost surely. It turns out that all the upper bounds can be easily obtained by an application of Inequality 5.1, combined with Inequality 5.5. Because of this similarity in proof for all the cases, we confine ourselves to giving a proof for part I.

By the Borel-Cantelli lemma and Inequality 5.5 it suffices to show that for every small positive ε we have $\sum PA_k < \infty$, where

$$(5.36) \quad A_k = \{\omega_{n_{k+1}}(a_{n_k}) \geq (2(1+2\varepsilon)(1-\varepsilon))^{-1} (1+c)a_{n_{k+1}} \log \log n_k)^{\frac{1}{2}}\}$$

and $n_k = [(1+\frac{1}{2}\varepsilon)^k]$. By an application of Inequality 5.1 we have

$$(5.37) \quad PA_k \leq C \frac{1}{a_{n_k}} \left(\log \frac{1}{a_{n_k}}\right)^{d-1} \exp(-(1+2\varepsilon)(1+c)(\log \log n_k) \frac{a_{n_{k+1}}}{a_{n_k}}) \\ \cdot \psi\left(\frac{(2(1+2\varepsilon)(1-\varepsilon))^{-1} (1+c)a_{n_{k+1}} a_{n_k}^{-1} \log \log n_k)^{\frac{1}{2}}}{(n_{k+1} a_{n_k})^{\frac{1}{2}}}\right).$$

By the conditions of part I it is easily seen that the argument of the function ψ tends to 0 as $k \rightarrow \infty$. Hence we have by (5.37) for large k

$$(5.38) \quad PA_k \leq C \frac{1}{a_{n_k}} \left(\log \frac{1}{a_{n_k}}\right)^{d-1} \exp(-(1+\varepsilon)(1+c) \log \log n_k) \\ = C \frac{1}{a_{n_k}} \left(\log \frac{1}{a_{n_k}}\right)^{d-1} \left(\frac{1}{\log n_k}\right)^{(1+\varepsilon)(1+c)}.$$

Using again the conditions of part I it follows from (5.38) that the PA_k are summable, which completes this part of the proof.

Let us next turn to the proof that these same numbers are lower bounds for the expressions on the left almost surely. Due to the fact that all the results in this theorem are independent of the dimension, we can use the onedimensional lower bounds whenever they are available. These lower bounds

can indeed be found in the literature for parts I, II, III and IV, namely in Mason, Shorack & Wellner (1983, Theorem 3), Stute (1982, Lemma 2.9), Komlós, Major & Tusnády (1975b, (3.11)) and Alexander (1984b, Theorem 3.1(C) and its proof) respectively. Hence it only remains to consider the lower bounds in part V.

We first consider the case $c/(c-1) \in \mathbb{N}$. Since $na_n \rightarrow 0$, it suffices to prove

$$(5.39) \quad \limsup_{k \rightarrow \infty} \sup_{R \in \bar{R}, |R| \leq a_{n_{k+1}}} (n_{k+1}^{-n_k}) F_{n_k, n_{k+1}} \{R\} \geq c/(c-1) \quad \text{a.s.,}$$

with $n_k = 2^k$ and $F_{n,m}$ defined as in (4.81). By straightforward computation we have for $m \in \mathbb{N}_0$ and a binomial rv $B(n,p)$ with $np < \frac{1}{2}$ and $n \geq 2m$

$$(5.40) \quad P(B(n,p) \geq m) \geq C(m)(np)^m, \quad C(m) \in (0,1).$$

Hence we have for large k , with \bar{R}_α as in the proof of Inequality 5.2,

$$\begin{aligned} (5.41) \quad & P(\sup_{|R| \leq a_{n_{k+1}}} (n_{k+1}^{-n_k}) F_{n_k, n_{k+1}} \{R\} \geq c/(c-1)) \\ & \geq P(\max_{R \in \bar{R}_{a_{n_{k+1}}}} (n_{k+1}^{-n_k}) F_{n_k, n_{k+1}} \{R\} \geq c/(c-1)) \\ & \geq 1 - \prod_{R \in \bar{R}_{a_{n_{k+1}}}} (1 - P((n_{k+1}^{-n_k}) F_{n_k, n_{k+1}} \{R\} \geq c/(c-1))) \\ & \geq 1 - \{1 - C(c/(c-1)) (2^k 2^{-(k+1)c})^{c/(c-1)}\}^{2^{(k+1)c}} \\ & \rightarrow 1 - \exp(-C(c/(c-1)) 2^{-c/(c-1)}) \text{ as } k \rightarrow \infty, \end{aligned}$$

where for the second inequality again Fact 4.3 is applied. Now (5.39) follows by an application of the Borel-Cantelli lemma.

Finally we consider the case $c/(c-1) \notin \mathbb{N}$. Then we have $\log(\frac{1}{a_n})/\log n \rightarrow c$, which is equivalent with $a_n = n^{-c+O(1)}$. Again since $na_n \rightarrow 0$, it suffices to prove

$$(5.42) \quad \liminf_{n \rightarrow \infty} \sup_{R \in \bar{R}; |R| \leq a_n} nF_n\{R\} \geq [c/(c-1)] \quad \text{a.s.}$$

So we have to show that

$$(5.43) \quad \sum_{n=1}^{\infty} P(\sup_{|R| \leq a_n} nF_n\{R\} \leq [c/(c-1)]-1) < \infty.$$

Defining $\mu \in (0,1)$ by $\mu = c/(c-1) - [c/(c-1)]$ and using again \bar{R}_α , Fact 4.3 and (5.40) we have for large n

$$\begin{aligned} (5.44) \quad & P(\sup_{|R| \leq a_n} nF_n\{R\} \leq [c/(c-1)]-1) \\ & \leq P(\max_{R \in \bar{R}_{a_n}} nF_n\{R\} \leq [c/(c-1)]-1) \\ & \leq \prod_{R \in \bar{R}_{a_n}} (1 - P(nF_n\{R\} \geq [c/(c-1)])) \\ & \leq \{1 - C([c/(c-1)]) n^{-(c-1)[c/(c-1)] + o(1)}\}^{n^{c+o(1)}} \\ & \leq \exp(-C([c/(c-1)]) n^{\frac{1}{2}(c-1)\mu}). \end{aligned}$$

These last numbers in (5.44) are summable since $\frac{1}{2}(c-1)\mu > 0$. This proves (5.42) and hence Theorem 5.3. □

3. Discussion and bibliography

The investigations which led to Theorems 5.1 and 5.2 were motivated by Silverman (1983), where the convergence to zero (as $n \rightarrow \infty$ and then $\alpha \downarrow 0$) of the expectation of the oscillation modulus of the *weighted* univariate uniform empirical process plays an important role in establishing weak convergence of weighted empirical processes of U-statistic structure. It should be noted that in these theorems the lower bounds for arbitrary $v \in \mathbb{N}$ can easily be derived from the case $v = 1$ and the Jensen inequality. A slightly weaker version of Theorem 5.2 is contained in Einmahl & Ruymgaart (1984).

Inequality 5.1 is a generalization in various respects of several

results in the literature. Univariate versions of this inequality are established in Stute (1982, Lemma 2.4) and Mason, Shorack & Wellner (1983, Inequality 1). A multivariate version of Inequality 5.1 is given in Stute (1984, Theorem 1.5). His inequality, however, is only applicable for relatively large sequences $\{a_n\}_{n=1}^{\infty}$ (see Theorem 5.3). It should also be noted that in his paper the upper bound for the exceedance probability of the oscillation modulus is essentially the same as that given in Inequality 5.1. His definition of the oscillation modulus is more restrictive than ours: conditions are imposed on the sides of the rectangles, which allow onedimensional techniques to be applied. A similar oscillation modulus will be considered further on in the discussion.

Section 2 of this chapter is contained in Einmahl & Ruymgaart (1986). Inequality 5.5 is an improvement of a maximal inequality which is given implicitly in the proof of Lemma 2.6 in Stute (1982) and explicitly in Mason, Shorack & Wellner (1983, Inequality 2), also compare our Inequality 2.11 (with $v = \frac{1}{2}$) with Stute's result. Inequality 5.5 generalizes both aforementioned onedimensional results to dimension $d \in \mathbb{N}$, and improves them as far as the conditions on α and λ are concerned. Moreover, in our upper bound the maximal index of the block appears, rather than the maximal index of the next block. In this way our maximal inequality resembles Kolmogorov type inequalities like Inequality 2.1.

Before discussing Theorem 5.3 we first define some oscillation moduli which are related to ω_n and derive some results for them. We define for $0 \leq \alpha \leq 1$:

$$(5.45) \quad \omega'_n(\alpha) = \sup_{R(s,t) \in \tilde{R}; \max_{1 \leq j \leq d} |t_j - s_j| \leq \alpha^{1/d}} |U_n\{R(s,t)\}|,$$

$$(5.46) \quad \bar{\omega}_n(\alpha) = \sup_{R \in \tilde{R}; |R| = \alpha} |U_n\{R\}|,$$

$$(5.47) \quad \Omega_n(\alpha) = \sup_{R \in \tilde{R}; |R| \leq \alpha} |U_n\{R\}|,$$

where \tilde{R} denotes the class of all closed rectangles in I^d . It is immediate that $\omega'_n(\alpha) \leq \omega_n(\alpha)$, hence it follows that the upper bounds in Theorem 5.3 are also upper bounds for ω'_n . Since in the case $d = 1$ the modulus ω'_n reduces to ω_n , the lower bounds of these moduli coincide as well. So we have shown that Theorem 5.3 holds true with ω_n replaced by ω'_n . We also

have immediately that $\bar{\omega}_n(\alpha) \leq \omega_n(\alpha)$. It turns out that all the lower bounds for ω_n can actually be based on $\bar{\omega}_n$. Hence Theorem 5.3 holds also true for $\bar{\omega}_n$. It is easy to see that

$$(5.48) \quad \lim_{n \rightarrow \infty} n^{\frac{1}{2}} \Omega_n(0) = d \quad \text{a.s.}$$

This simple result shows that the almost sure behaviour of Ω_n depends on the dimension d , whereas the results for ω_n are independent of the dimension. This difference in the almost sure behaviour of the moduli Ω_n and ω_n not only shows itself for the degenerate sequence $\{0\}_{n=1}^{\infty}$. Consider a sequence of part V with $c > 2$; we have $\lim_{n \rightarrow \infty} n^{\frac{1}{2}} \omega_n(a_n) = 1$ a.s., whereas it follows from (5.48) that $\liminf_{n \rightarrow \infty} n^{\frac{1}{2}} \Omega_n(a_n) \geq d$ a.s.

Theorem 5.3 generalizes and improves upon various results in the literature and gives a complete description of the almost sure behaviour of the limsup of properly standardized versions of ω_n . The onedimensional version of part I is established in Mason, Shorack & Wellner (1983), with the aid of a result in Chan (1977) and strong approximation techniques. Part II, for $d = 1$, is contained in Stute (1982) and for ω'_n and arbitrary $d \in \mathbb{N}$ in Stute (1984). Part III is proved for $d = 1$ in Mason, Shorack & Wellner (1983), where also a partial result for part IV ($d = 1$) is given. For ω'_n ($d \in \mathbb{N}$) part III is contained in Einmahl & Ruymgaart (1985a); for ω_n this part is established in Einmahl (1984b), along with an alternative proof. The sequences of part V so far have not been considered in the literature. It is pointless to consider smaller sequences than those in part V because the lim is almost surely equal to 1 if $c > 2$ and even $\lim_{n \rightarrow \infty} n^{\frac{1}{2}} \omega_n(0) = 1$.

CHAPTER 6

GENERALIZATIONS AND APPLICATIONS

This final chapter has a somewhat different character than the preceding ones. It consists of four more or less unconnected sections; in the first three sections generalizations of the results in Chapters 2-5 will be briefly considered, whereas in the last section some (non-standard) applications of the theory are mentioned. Section 1 will be devoted to the "negative parts" of empirical processes and in section 2 df's other than the uniform will be considered. A few miscellaneous results are discussed in section 3. It is beyond the scope of this monograph to present the proofs of the theorems in sections 1 and 2. Although occasionally there are subtle differences in detail, in principle their proofs are patterned on those of the corresponding theorems in the previous chapters.

1. The negative part of the empirical process

Let us define U_n^+ and U_n^- , the positive and negative parts of U_n , by

$$(6.1) \quad U_n^\pm(t) = \pm U_n(t) \vee 0, \quad t \in I^d,$$

and ω_n^+ and ω_n^- by

$$(6.2) \quad \omega_n^\pm(\alpha) = \sup_{R \in \mathcal{R}; |R| \leq \alpha} U_n^\pm\{R\}, \quad 0 \leq \alpha \leq 1.$$

By inspection of the proofs of the results in Chapters 3-5 for the absolute value of U_n and ω_n , it is easily seen that those results remain true for U_n^+ and ω_n^+ . Due to the asymmetry of the binomial distribution the situation is completely different for U_n^- (and ω_n^-). We will consider the latter process briefly in this section, presenting some of the most interesting

analogues for U_n^- of the results in Chapters 2-5. Surprisingly, we will see that U_n^- behaves much "better" than U_n^+ for relatively small indexing sets. In particular, truncation, as in Theorems 3.3 and 4.3, is superfluous now: we can take the suprema over \bar{R} and over the whole unit square respectively.

We begin with the analogue for U_n^- of Inequality 2.6; its proof will be omitted since it can be patterned on that of Inequality 2.6, using (2.18) instead of (2.19). We will replace $\psi(\cdot)(\geq 1)$ by 1, since it is not needed for most applications. Note that (2.18) is true for *all* $\lambda \geq 0$ if $\psi(-(1-\varepsilon)\lambda/(n^{\frac{1}{2}}|R|))$ is omitted. "Negative versions" of Inequalities 2.7-2.9 can be derived in the same way and will therefore not be stated here.

INEQUALITY 6.1. Let $q \in Q$. Then we have for any $\varepsilon \in (0,1)$ and $0 < \alpha \leq \beta \leq \frac{1}{2}(1-\varepsilon)$

$$(6.3) \quad P(\sup_{\alpha \leq |t| \leq \beta} U_n^-(t)/q(|t|) \geq \lambda) \\ \leq C \int_{(1-\varepsilon)\alpha}^{\beta/(1-\varepsilon)} \frac{(\log(1/\sigma))^{d-1}}{\sigma} \exp\left(\frac{-(1-\varepsilon)\lambda^2 q^2(\sigma)}{2\sigma}\right) d\sigma, \quad \lambda \geq 0,$$

where $C = C(d, \varepsilon) \in (0, \infty)$.

Now we turn to ω_n^- . The proof of the next inequality is immediate from the proof of Inequality 5.1.

INEQUALITY 6.2. Let $\varepsilon \in (0, \frac{1}{2}]$ and $\alpha \in (0, \frac{1}{4})$. Then we have

$$(6.4) \quad P(\omega_n^-(\alpha) \geq \lambda \alpha^{\frac{1}{2}}) \leq C \frac{1}{\alpha} (\log \frac{1}{\alpha})^{d-1} \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2} \psi\left(\frac{-(1-\varepsilon)\lambda}{(n\alpha)^{\frac{1}{2}}}\right)\right), \quad 0 \leq \lambda \leq (n\alpha)^{\frac{1}{2}},$$

where $C = C(d, \varepsilon) \in (0, \infty)$.

A brief examination of the proofs of the maximal inequalities presented in Chapters 2, 4 and 5, shows that all those inequalities remain valid for the corresponding negative (and positive) parts.

We now present the main results of this section followed, in conclusion, by a discussion.

THEOREM 6.1. Let $d \in \mathbb{N}$ and $q \in \mathcal{Q}$. Then we have in the setup of Chapter 3

$$(6.5) \quad \sup_{t \in I^d} |U_n^-(t) - U^-(t)| / q(|t|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if $q \in \mathcal{Q}_0$, where $U^-(t) = -U(t) \vee 0$, $t \in I^d$.

THEOREM 6.2. Let $d \in \mathbb{N}$ and $q \in \mathcal{Q}^*$. Then we have in the setup of Chapter 3

$$(6.6) \quad \sup_{R \in \mathbb{R}} |U_n^-\{R\} - U^-\{R\}| / q(|R|) \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

if and only if $q \in \mathcal{Q}_1$.

THEOREM 6.3. Let $d \in \mathbb{N}$. Then we have

$$(6.7) \quad \limsup_{n \rightarrow \infty} \sup_{t \in I^d} \frac{U_n^-(t)}{(\log \log n)^{\frac{1}{2}} |t|^{\frac{1}{2}}} = (2(d+1))^{\frac{1}{2}} \text{ a.s.}$$

THEOREM 6.4. Let $d \in \mathbb{N}$ and $0 \leq \nu < \frac{1}{2}$. Then we have in the obvious notation

$$(6.8) \quad \limsup_{n \rightarrow \infty} \sup_{t \in I^d} \frac{U_{n,\nu}^-(t)}{(\log \log n)^\nu |t|^{1-\nu}} = \max_{c \geq d} c^\nu (1 - \beta_{c/d}^-) \text{ a.s.,}$$

where, for $\sigma \in (1, \infty)$, β_σ^- is the solution in $(0, 1)$ to the equation in (4.62) and, for $\sigma \in (0, 1]$, $\beta_\sigma^- = 0$.

THEOREM 6.5. Let $d \in \mathbb{N}$ and $\{a_n\}_{n=1}^\infty$ as in (4.61). Then we have

$$(6.9) \quad \limsup_{n \rightarrow \infty} \sup_{a_n \leq |t| \leq 1} |t| / F_n(t) = 1 / \beta_{c/d}^- \text{ a.s.}$$

THEOREM 6.6. Let $d \in \mathbb{N}$.

(I) If $\log(1/a_n) / \log \log n \rightarrow c \in [0, \infty)$ and $na_n \uparrow$ if $c = 0$, then

$$(6.10) \quad \limsup_{n \rightarrow \infty} \frac{\omega_n^-(a_n)}{(a_n \log \log n)^{\frac{1}{2}}} = (2(1+c))^{\frac{1}{2}} \text{ a.s.}$$

(II) If $\log(1/a_n) / \log \log n \rightarrow \infty$, $na_n / \log n \rightarrow \infty$ and $na_n \uparrow$, then

$$(6.11) \quad \lim_{n \rightarrow \infty} \frac{\omega_n^-(a_n)}{(a_n \log(1/a_n))^{\frac{1}{2}}} = 2^{\frac{1}{2}} \quad \text{a.s.}$$

(III) If $na_n/\log n \rightarrow c \in [0, \infty)$, then

$$(6.12) \quad \lim_{n \rightarrow \infty} \frac{\omega_n^-(a_n)}{n^{\frac{1}{2}} a_n} = 1 - \beta_c^- \quad \text{a.s.}$$

Inequality 6.1 is the generalization to dimension d of the negative part of Inequality 1.1 in Shorack & Wellner (1982), whereas Inequality 6.2 generalizes the corresponding part of Inequality 1 in Mason, Shorack & Wellner (1983) w.r.t. the dimension. Note, however, that they omit the function ψ in their upper bound, which leads to the incorrect value for the limit in their statement (6-) for $c > 1$. Our Theorem 6.6, part III, corrects their statement (6-) in this case. Theorem 6.3 generalizes Theorem 3.4 in Csáki (1977) to dimension d , cf. also Theorem 1 in Shorack (1980). Since $\max_{c \geq d} c^v (1 - \beta_{c/d}^-) < (2(d+1))^{\frac{1}{2}}$, for $0 \leq v < \frac{1}{2}$, in contrast to Theorem 4.3, it is more convenient to state the results separately in (6.7) and (6.8). Theorem 6.5 generalizes Theorem 2(ii) in Wellner (1978). We included it in this section, since for $\alpha \in (0, 1)$ and $\lambda \geq 1$

$$(6.13) \quad \left\{ \sup_{\alpha \leq |t| \leq 1} |t|/F_n(t) \geq \lambda \right\} = \left\{ \sup_{\alpha \leq |t| \leq 1} U_{n,0}^-(t)/|t| \geq 1 - 1/\lambda \right\}.$$

Finally, we remark that the analogues for the negative parts of the results in Chapters 3-5 which are not presented in this section, can also be derived by our tools and techniques. For the sake of brevity, we do not present them here.

2. Extensions to a continuous density bounded away from 0 and ∞

In this section, the underlying df is not necessarily the uniform one. Instead it will be assumed more generally for the df F of the X_i , defined in Chapter 1, that

$$(6.14) \quad \begin{cases} F \text{ has a continuous density } f \text{ w.r.t. Lebesgue measure, such that} \\ 0 < m = \inf_{t \in I^d} f(t) \leq \sup_{t \in I^d} f(t) = M < \infty. \end{cases}$$

In this setup we obtain in a number of cases analogues for U_n^F , see (1.3), of the results in Chapters 2-5. Although assumption (6.14) may be weakened somewhat in the statements of the results in this section, we recall that e.g. Theorem 3.1 cannot hold for all df's F , in particular nor for \bar{F} (see (3.38) and (2.34)). We restrict ourselves, for the sake of convenience and brevity, to the empirical process indexed by points. The process indexed by rectangles, however, can be treated along similar lines. We make one exception in the case of the indexing class \bar{R} and present the analogue for ω_n^F (defined below) of Theorem 5.3.

As in section 1 we begin with the analogue of Inequality 2.6; the analogue of Inequality 2.8 can be derived in a similar way.

INEQUALITY 6.3. Let $q \in \mathcal{Q}$. Then we have for any $\varepsilon \in (0,1)$ and $0 < \alpha \leq \beta \leq \frac{1}{2}(1-\varepsilon)$

$$(6.15) \quad P\left(\sup_{\alpha \leq F(t) \leq \beta} |U_n^F(t)|/q(F(t)) \geq \lambda\right) \\ \leq C \int_{(1-\varepsilon)\alpha}^{\beta/(1-\varepsilon)} \frac{(\log(1/\sigma))^{d-1}}{\sigma} \exp\left(-\frac{(1-\varepsilon)\lambda^2 q^2(\sigma)}{2\sigma}\right) \psi\left(\frac{\lambda q(\alpha)}{n^{\frac{1}{2}}\alpha}\right) d\sigma, \quad \lambda \geq 0,$$

where $C = C(d, \varepsilon, m, M) \in (0, \infty)$.

PROOF. The proof is along similar lines to that of Inequality 2.6. Therefore we will restrict ourselves to the two main differences that appear, namely the analogue of (2.23) and the change of variables below (2.31). It is easily seen that instead of (2.23) it suffices to show that for any $\varepsilon \in (0,1)$ there exists a $\theta \in (0,1)$ such that for any $R(a,b) \in \mathcal{P}(\theta)$

$$(6.16) \quad \frac{F(b)}{F(a)} \geq 1-\varepsilon.$$

Let us write $E = R(\langle 0 \rangle, b) \setminus R(\langle 0 \rangle, a)$. Then we have in the obvious notation

$$(6.17) \quad \frac{F(b)}{F(a)} = 1 + \frac{F\{E\}}{F(a)} \leq 1 + \frac{M|E|}{m|a|} = 1 + \frac{M(1-\theta^d)}{m\theta^d}.$$

Now, by letting $\theta \uparrow 1$, it is easily seen that (6.16) holds true.

Replace the change of variables below (2.31) by $\sigma = s_1 = F(t)$,

$s_2 = t_2, \dots, s_d = t_d$. Then elementary analysis, using (6.14), easily completes the proof. □

Let us now define ω_n^F by

$$(6.18) \quad \omega_n^F(\alpha) = \sup_{R \in \mathcal{R}; |R| \leq \alpha} |U_n^F\{R\}|, \quad 0 \leq \alpha \leq 1;$$

note that for an oscillation modulus we have to consider rectangles with $|R| \leq \alpha$, rather than $F\{R\} \leq \alpha$, cf. the left side of (6.15). Following the proof of Inequality 5.1 we obtain

INEQUALITY 6.4. Let $\varepsilon \in (0, \frac{1}{2}]$ and $\alpha \in (0, 1/(4M))$. Then we have

$$(6.19) \quad P(\omega_n^F(\alpha) \geq \lambda(\alpha M)^{\frac{1}{2}}) \leq C \frac{1}{\alpha} (\log \frac{1}{\alpha})^{d-1} \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2} \psi\left(\frac{\lambda}{(n\alpha M)^{\frac{1}{2}}}\right)\right), \quad \lambda \geq 0,$$

where $C = C(d, \varepsilon) \in (0, \infty)$.

Observe that also all the appropriate maximal inequalities remain true, mutatis mutandis, under assumption (6.14), as indicated above Inequality 2.11.

We now present our results. Since the statements of most of the theorems do not change drastically when generalized to the non-uniform case (6.14), we will only state two such theorems explicitly, namely the analogues of Theorems 4.1 and 5.3. The other theorems will be only discussed. For the analogue of Theorem 4.1 we need the following definition:

$$(6.20) \quad V_{n,v}^F = \sup_{t \in I^d} \frac{n^v |F_n(t) - F(t)|}{(F(t)(1-F(t)))^{1-v}}, \quad 0 \leq v \leq \frac{1}{2}.$$

THEOREM 6.7. Let $d \in \mathbb{N}$, $0 \leq v \leq \frac{1}{2}$ and $\{a_n\}_{n=1}^{\infty}$ a sequence of positive constants.

(I) If $\sum_{n=1}^{\infty} a_n (\log(1/a_n))^{d-1} = \infty$, then

$$(6.21) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-v} V_{n,v}^F = \infty \quad \text{a.s.}$$

(II) If $\sum_{n=1}^{\infty} a_n (\log (1/a_n))^{d-1} < \infty$ and $na_n \downarrow$, then

$$(6.22) \quad \lim_{n \rightarrow \infty} (na_n)^{1-\nu} V_{n,\nu}^F = 0 \quad \text{a.s.}$$

THEOREM 6.8. Apart from the numbers on the right, Theorem 5.3 remains exactly the same for ω_n^F . These numbers become for the respective parts:

$$(6.23) \quad (2(1+c)M)^{\frac{1}{2}}; (2M)^{\frac{1}{2}}; cM(\beta_{cM}-1); 1; [c/(c-1)].$$

Theorem 6.7 is implicitly contained in Einmahl & Mason (1985a). "Part II" of Theorem 6.8 is proved for $d = 1$ in Stute (1982) and for $d \in \mathbb{N}$, based on a restricted definition of the oscillation modulus (the generalization of ω_n' w.r.t. the df), in Stute (1984), where also a weaker version of Inequality 6.4 for that modulus is obtained. Using i.a. inequalities like Inequality 6.3 it can be seen that Theorems 3.1, 3.2, 4.2 and 5.2 remain also true, mutatis mutandis, under assumption (6.14). Replacing U_n by U_n^F and $|t|$ by $F(t)$ in Theorems 4.3-4.5 does not affect the upper bounds. It is likely that the same lower bounds can be derived by modifications of techniques like those used for the proof of (4.75). Also the required lower bounds may be contained implicitly in Alexander (1984b). Theorems 4.6 and 4.7 appear to remain unaltered under (6.14). It is interesting to observe that the replacement of ω_n by ω_n^F affects the numbers on the right in Theorem 5.3. This is due to the fact that $|R|$ (instead of $F\{R\}$) is maintained in the definition of ω_n^F .

3. Miscellaneous results

a. Empirical processes based on spacings. First we present the analogue of Theorem 4.1 for the empirical process based on uniform one-dimensional spacings. This result is stated and proved in Einmahl & van Zuijlen (1985); large parts of the proof are obtained by modifications of the corresponding parts of the proof of Theorem 4.1. Let $\{X_i\}_{i=1}^{\infty}$ be the sequence of rv's defined in Chapter 1 ($d = 1$), with F the uniform df. For $n \geq 2$ define \tilde{F}_n to be the empirical df based on the transformed uniform spacings $\mathbb{D}_{j,n} = 1 - \exp(-nD_{j,n})$, $j = 1, \dots, n$, where $D_{j,n}$ is the j -th spacing, defined by $D_{j,n} = X_{j:n-1} - X_{j-1:n-1}$ ($X_{0:n-1} = 0$, $X_{n:n-1} = 1$). Notice that

the empirical process $n^{\frac{1}{2}}(\tilde{F}_n(t)-t)$, $t \in I$, is not included in the setup of this monograph since the rv's $D_{j,n}$ are dependent (and form a triangular array instead of a sequence). Finally define for $0 \leq v \leq \frac{1}{2}$

$$(6.24) \quad \tilde{V}_{n,v,\ell} = \sup_{t \in I} \frac{n^v |\tilde{F}_n(t)-t|}{t^{1-v}},$$

$$(6.25) \quad \tilde{V}_{n,v,r} = \sup_{t \in I} \frac{n^v |F_n(t)-t|}{(1-t)^{1-v}},$$

and note that ℓ or r indicates whether the process is weighted in the left or right tail.

THEOREM 6.9. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of positive constants and let $0 \leq v \leq \frac{1}{2}$.

(I ℓ) If $\sum_{n=1}^{\infty} a_n = \infty$, then

$$(6.26) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-v} \tilde{V}_{n,v,\ell} = \infty \quad \text{a.s.}$$

(II ℓ) If $\sum_{n=1}^{\infty} a_n < \infty$, $a_n \downarrow$ and $na_n \log n \rightarrow 0$, then

$$(6.27) \quad \lim_{n \rightarrow \infty} (na_n)^{1-v} \tilde{V}_{n,v,\ell} = 0 \quad \text{a.s.}$$

(I r) If $\sum_{n=1}^{\infty} a_n \log(1/a_n) = \infty$, then

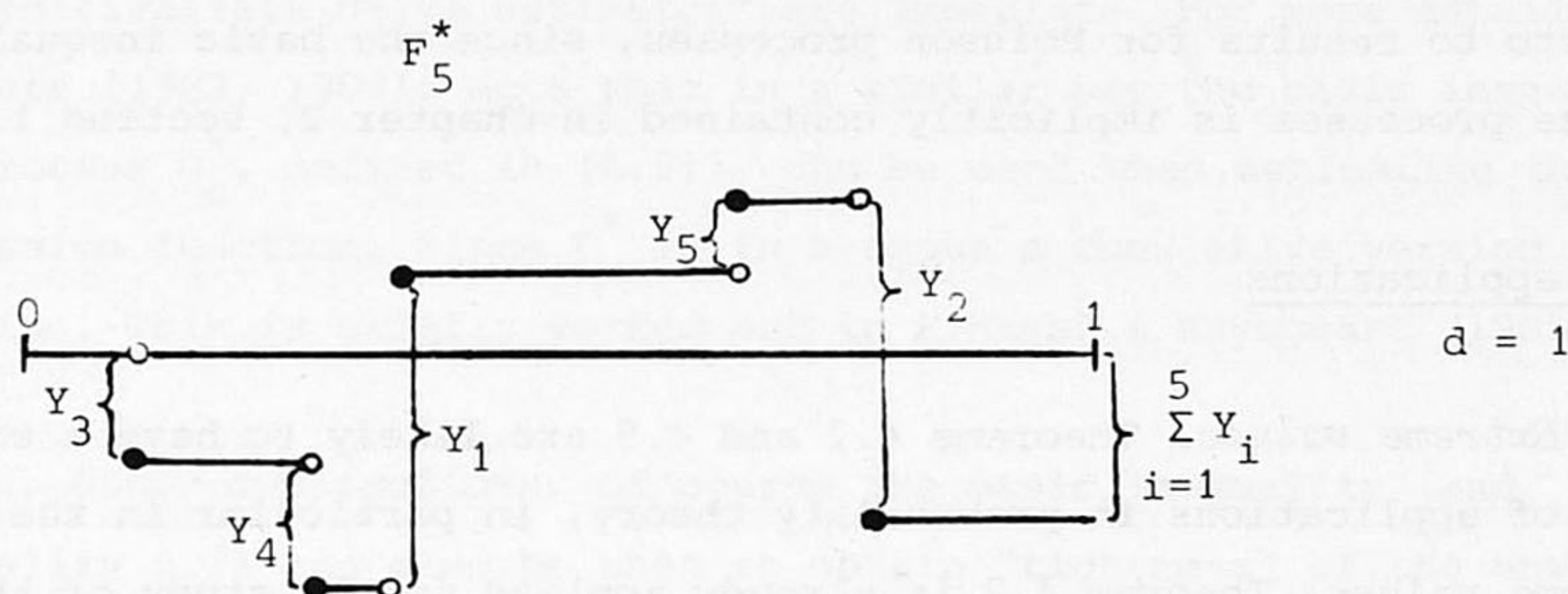
$$(6.28) \quad \limsup_{n \rightarrow \infty} (na_n)^{1-v} \tilde{V}_{n,v,r} = \infty \quad \text{a.s.}$$

(II r) If $\sum_{n=1}^{\infty} a_n \log(1/a_n) < \infty$ and $a_n \downarrow 0$, then

$$(6.29) \quad \lim_{n \rightarrow \infty} (na_n)^{1-v} \tilde{V}_{n,v,r} = 0 \quad \text{a.s.}$$

b. Compound empirical processes. Another situation which can not be covered in our general format occurs when the jumps of the empirical df are allowed to be rv's. We will briefly explain and discuss this case here. Let $\{<X_i, Y_i>\}_{i=1}^{\infty}$ be a sequence of iid random vectors, where the X_i take their values in I^d and the Y_i in a bounded subset $A \subset \mathbb{R}$. Typically X_i and Y_i are independent. Now define the so called compound empirical df by

$$(6.30) \quad F_n^*(t) = \frac{1}{n} \sum_{i=1}^n Y_i 1_{R(<0>,t)}(X_i), \quad t \in I^d \quad (R(<0>,t) \in \bar{R}),$$



and the compound multivariate empirical process by

$$(6.31) \quad U_n^*(t) = n^{\frac{1}{2}} (F_n^*(t) - F^*(t)), \quad t \in I^d,$$

where $F^*(t) = EY_i 1_{R(<0>,t)}(X_i)$. Observe that $U_n^*(t)$ reduces to the ordinary empirical process when the distribution of the Y_i is degenerate at 1. In order to define a proper weighting for the process U_n^* , we set

$$(6.32) \quad V(t) = EY_i^2 1_{R(<0>,t)}(X_i), \quad t \in I^d.$$

Now, with modifications of the techniques used in the previous chapters, in particular by relating U_n^* to a *compound* Poisson process, it is possible, under some additional assumptions, to obtain asymptotic results for the weighted compound multivariate empirical process

$$(6.33) \quad U_n^*(t)/q(V(t)), \quad t \in I^d \quad (q \in Q^*).$$

For more details see Einmahl & Ruymgaart (1985b), where in particular a weak convergence result and a strong limit theorem for the process in (6.33) are obtained. Also refer to Marcus & Zinn (1984).

c. Extension to \mathbb{R}^d . It is easily seen that the theorems in this monograph extend to empirical processes on \mathbb{R}^d by applying suitable trans-

formations ϕ_j , say, on each coordinate t_j of $t \in (0,1)^d$, $1 \leq j \leq d$, i.e. a transformation $\phi : (0,1)^d \rightarrow \mathbb{R}^d$, with $\phi(t) = \langle \phi_j(t_j) \rangle$, $t \in I^d$.

d. Poisson processes. It is also obvious that our approach will lead ipso facto to results for Poisson processes, since the basic inequality for these processes is implicitly contained in Chapter 2, section 1.

4. Some applications

a. Extreme values. Theorems 4.2 and 4.5 are likely to have a wide variety of applications in probability theory, in particular in the theory of extreme values. Theorem 4.2 is already applied in the study of the almost sure stability of sums of extreme values in Einmahl, Haeusler & Mason (1985). Theorem 4.5, in turn, has already proven to be a valuable tool in establishing LIL's for sums of extreme values, see Haeusler & Mason (1985) and Deheuvels, Haeusler & Mason (1985).

b. U-statistics. A nice application of the basic inequality, can be found in Helmers & Ruymgaart (1986), where the inequality in (2.18) is applied to obtain a similar inequality for empirical processes of U-statistic structure. Their inequality is derived using a moment generating function technique; statement (5.3) along with (2.18) is then used to bound the moment generating function from above.

c. General spacings. Let us define $\bar{\omega}_n^\pm$ in the obvious way, see (5.46) and (6.2). Then it is straightforward that Theorems 5.3 and 6.6 remain true for $\bar{\omega}_n^+$ and $\bar{\omega}_n^-$ respectively. Combination of the thus obtained results and a slight modification of an observation in Mason (1984) leads to theorems on the almost sure behaviour of properly defined multivariate maximal or minimal k -spacings, where k depends on n . It should be noted that these spacings are defined in terms of rectangles, but it is also possible to use a more general setup, where the class of rectangles is replaced by a class of sets satisfying certain entropy conditions. For a detailed study of these spacings in the latter setup the reader is referred to Deheuvels, Einmahl, Mason & Ruymgaart (1986).

d. Density and regression function estimation. With the aid of suitable transformations the basic inequality has also applications in the theory of density estimation on \mathbb{R}^d . Especially strong convergence properties of the "multivariate naive estimator" are immediate. For more details we refer to Stute (1982, 1984). Note that in a similar way the basic inequality for the process U_n^* , defined in (6.31), can be used when estimating the multivariate regression function, since F^* is in a sense a cumulative version of this function. This is briefly worked out in Einmahl & Ruymgaart (1985b).

e. Other applications. Of course the basic inequality (and Inequality 5.3) can also be used to obtain "tightness" of the unweighted empirical processes, i.e. that our theory also provides a proof for Fact 1.1.A. For some other applications of the theory of empirical processes (pattern recognition, cluster analysis, etc.) we refer to Dudley (1982), Pollard (1981, 1982) and Stute (1982, 1984).

REFERENCES

- [1] K.S. Alexander (1982). *Some Limit Theorems for Weighted and Non-Identically Distributed Empirical Processes*, Ph.D. Dissertation, M.I.T., Cambridge.
- [2] K.S. Alexander (1984a). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* 12, 1041-1067.
- [3] K.S. Alexander (1984b). *Rates of Growth and Sample Moduli for Weighted Empirical Processes Indexed by Sets*, Technical Report, University of Washington, Seattle.
- [4] K.S. Alexander (1984c). *The Central Limit Theorem for Empirical Processes on Vapnik-Červonenkis Classes*, Technical Report, University of Washington, Seattle.
- [5] K.S. Alexander (1984d). Personal communication.
- [6] K.S. Alexander (1985). *The Central Limit Theorem for Weighted Empirical Processes Indexed by Sets*, Technical Report, University of Washington, Seattle.
- [7] O. Barndorff-Nielsen (1961). On the rate of growth of the partial maxima of a sequence of independent identically distributed random variables. *Math. Scand.* 9, 383-394.
- [8] P. Billingsley (1968). *Convergence of Probability Measures*, Wiley, New York.
- [9] A.H.C. Chan (1977). *On the Increments of Multiparameter Gaussian Processes*, Ph.D. Dissertation, Carleton University, Ottawa.
- [10] D.M. Chibisov (1964). Some theorems on the limiting behaviour of the empirical distribution function. *Sel. Transl. Math. Statist. Prob.* 6, 147-156.
- [11] K.L. Chung (1949). An estimate concerning the Kolmogorov limit distribution. *Trans. Amer. Math. Soc.* 67, 36-50.
- [12] E. Csáki (1974). Studies on the empirical d.f. *MTA III. Oszt. Közl.* 23, 239-327 (in Hungarian).
- [13] E. Csáki (1975). Some notes on the law of the iterated logarithm for empirical distribution function, P. Révész (ed.). *Colloq. Math. Soc. János Bolyai* 11, North-Holland, Amsterdam, 47-58.

- [14] E. Csáki (1977). The law of the iterated logarithm for normalized empirical distribution function. *Z. Wahrsch. Verw. Gebiete* 38, 147-167.
- [15] E. Csáki (1982). On the standardized empirical distribution function. B.V. Gnedenko, M.L. Puri, I. Vincze (eds.). *Colloq. Math. Soc. János Bolyai* 32, North-Holland, Amsterdam, 123-138.
- [16] M. Csörgö, S. Csörgö, L. Horváth, D.M. Mason (1983). *An Asymptotic Theory for Empirical Reliability and Concentration Processes*, Technical Report No. 2, Laboratory for Research in Statistics and Probability, Ottawa.
- [17] M. Csörgö, S. Csörgö, L. Horváth, D.M. Mason (1986). Weighted empirical and quantile processes. *Ann. Probab.* 14, 31-85.
- [18] M. Csörgö, D.M. Mason (1985). On the asymptotic distribution of weighted empirical and quantile processes in the middle and on the tails. *Stochastic Process Appl.* 21, 119-132.
- [19] M. Csörgö, P. Révész (1981). *Strong Approximations in Probability and Statistics*, Academic Press, New York.
- [20] P. Deheuvels, J.H.J. Einmahl, D.M. Mason, F.H. Ruymgaart (1986). *The Almost Sure Behaviour of Maximal and Minimal Multivariate k_n -Spacings*, in preparation.
- [21] P. Deheuvels, E. Haeusler, D.M. Mason (1985). *On the Almost Sure Behaviour of Sums of Extreme Values from a Distribution in the Domain of Attraction of a Gumbel Law*, Preprint.
- [22] L. Devroye (1982). Bounds for the uniform deviation of empirical measures, *J. Multivariate Anal.* 12, 72-79.
- [23] M.D. Donsker (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* 23, 277-281.
- [24] R.M. Dudley (1978). Central limit theorems for empirical measures. *Ann. Probab.* 6, 899-929.
- [25] R.M. Dudley (1982). Some recent results on empirical processes. A. Beck (ed.). *Probability in Banach Spaces III; Lecture Notes in Mathematics* 860, Springer-Verlag, New York, 107-123.
- [26] R.M. Dudley (1984). A course on empirical processes. P.L. Hennequin (ed.). *École d'Été de Probabilités de Saint-Flour XII-1982; Lecture Notes in Mathematics* 1097, Springer-Verlag, New York, 1-142.
- [27] F. Eicker (1979). The asymptotic distribution of the suprema of the standardized empirical process. *Ann. Statist.* 7, 116-138.

- [28] J.H.J. Einmahl (1984a). *A General Form of the Law of the Iterated Logarithm for the Weighted Multivariate Empirical Distribution Function*, Report 8418, Catholic University, Nijmegen.
- [29] J.H.J. Einmahl (1984b). *A Strong Law for the Oscillation Modulus of the Multivariate Empirical Process II*, Report 8422, Catholic University, Nijmegen.
- [30] J.H.J. Einmahl, E. Haeusler, D.M. Mason (1985). *An Extension of a Theorem of Csáki with Application to the Study of the Almost Sure Stability of Sums of Extreme Values*, Report 8537, Catholic University, Nijmegen.
- [31] J.H.J. Einmahl, D.M. Mason (1985a). Bounds for weighted multivariate empirical distribution functions. *Z. Wahrsch. Verw. Gebiete* 70, 563-571.
- [32] J.H.J. Einmahl, D.M. Mason (1985b). *Laws of the Iterated Logarithm in the Tails for Weighted Uniform Empirical Processes*, Report 8534, Catholic University, Nijmegen.
- [33] J.H.J. Einmahl, F.H. Ruymgaart (1984). *The Order of Magnitude of the Moments of the Modulus of Continuity of Multiparameter Poisson and Empirical Processes*, Report 8402, Catholic University, Nijmegen.
- [34] J.H.J. Einmahl, F.H. Ruymgaart (1985a). A strong law for the oscillation modulus of the multivariate empirical process. *Statist. Decisions* 3, 357-362.
- [35] J.H.J. Einmahl, F.H. Ruymgaart (1985b). Some properties of weighted compound multivariate empirical processes. *Sankhyā Ser. A*, to appear.
- [36] J.H.J. Einmahl, F.H. Ruymgaart (1986). *The Almost Sure Behaviour of the Oscillation Modulus of the Multivariate Empirical Process*, Report 8611, Catholic University, Nijmegen.
- [37] J.H.J. Einmahl, F.H. Ruymgaart, J.A. Wellner (1984). A characterization of weak convergence of weighted multivariate empirical processes. *Acta Sci. Math. (Szeged)*, to appear.
- [38] J.H.J. Einmahl, M.C.A. van Zuijlen (1985). *Strong Bounds for Weighted Empirical Distribution Functions Based on Uniform Spacings*, Report 8531, Catholic University, Nijmegen.
- [39] P. Gaenssler (1983). *Empirical Processes*, IMS Lecture Notes - Monographic Series 3, Hayward.
- [40] P. Gaenssler, W. Stute (1979). Empirical Processes: a survey of results for independent and identically distributed random variables. *Ann. Probab.* 7, 193-243.

- [41] J. Geffroy (1958/1959). Contributions à la théorie des valeurs extrêmes. *Publ. Inst. Statist. Univ. Paris* 7/8, 37-185.
- [42] E. Haeusler, D.M. Mason (1985). A law of the iterated logarithm for sums of extreme values from a distribution with a regularly varying upper tail. *Ann. Probab.*, to appear.
- [43] R. Helmers, F.H. Ruymgaart (1986). *Asymptotic Normality of GL-Statistics with Unbounded Scores*, in preparation.
- [44] D. Jaeschke (1979). The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. *Ann. Statist.* 7, 108-115.
- [45] B.R. James (1975). A functional law of the iterated logarithm for weighted empirical distributions. *Ann. Probab.* 3, 762-772.
- [46] J. Kiefer (1961). On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* 11, 649-660.
- [47] J. Kiefer (1972). Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. L.M. LeCam, J. Neyman, E.L. Scott (eds.). *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability I*, University of California Press, Berkeley, 227-244.
- [48] J. Kiefer, J. Wolfowitz (1958). On the deviations of the empiric distribution function of vector chance variables. *Trans. Amer. Math. Soc.* 87, 173-186.
- [49] O.I. Klesov (1983). The law of the iterated logarithm for multiple sums. *Theory Probab. Math. Statist.* 27, 65-72.
- [50] A. Kolmogorov (1929). Uber das Gesetz des iterierten Logarithmus. *Math. Ann.* 101, 126-135.
- [51] J. Komlós, P. Major, G. Tusnády (1975a). An approximation of partial sums of independent rv's, and the sample df, I. *Z. Wahrsch. Verw. Gebiete* 32, 111-131.
- [52] J. Komlós, P. Major, G. Tusnády (1975b). Weak convergence and embedding. P. Révész (ed.). *Colloq. Math. Soc. János Bolyai* 11, North-Holland, Amsterdam, 149-165.
- [53] M. Loève (1977). *Probability Theory I* (4th. ed.), Springer-Verlag, New York.
- [54] C.L. Mallows (1968). An inequality involving multinomial probabilities. *Biometrika* 55, 422-424.
- [55] M.B. Marcus, J. Zinn (1984). The bounded law of the iterated logarithm for the weighted empirical distribution process in the non-i.i.d. case. *Ann. Probab.* 12, 335-360.

- [56] D.M. Mason (1981). Bounds for weighted empirical distribution functions. *Ann. Probab.* 9, 881-884.
- [57] D.M. Mason (1982). Some characterizations of almost sure bounds for weighted multidimensional empirical distributions and a Glivenko-Cantelli theorem for sample quantiles. *Z. Wahrsch. Verw. Gebiete* 59, 505-513.
- [58] D.M. Mason (1983). The asymptotic distribution of weighted empirical distribution functions. *Stochastic Process. Appl.* 15, 99-109.
- [59] D.M. Mason (1984). A strong limit theorem for the oscillation modulus of the uniform empirical quantile process. *Stochastic Process. Appl.* 17, 127-136.
- [60] D.M. Mason (1985). The asymptotic distribution of generalized Rényi statistics. *Acta Sci. Math. (Szeged)* 48, 315-323.
- [61] D.M. Mason, G.R. Shorack, J.A. Wellner (1983). Strong limit theorems for oscillation moduli of the uniform empirical process. *Z. Wahrsch. Verw. Gebiete* 65, 83-97.
- [62] D.M. Mason, W.R. van Zwet (1985). *A Refinement of the KMT Inequality for the Uniform Empirical Process*, Preprint No. 30, University of Munich, Munich.
- [63] D.S. Moore, M.C. Spruill (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *Ann. Statist.* 3, 599-616.
- [64] T. Mori (1976). The strong law of large numbers when extreme terms are excluded from sums. *Z. Wahrsch. Verw. Gebiete* 36, 189-194.
- [65] G. Neuhaus (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* 42, 1285-1295.
- [66] N.E. O'Reilly (1974). On the weak convergence of empirical processes in sup-norm metrics. *Ann. Probab.* 2, 642-651.
- [67] S. Orey, W.E. Pruitt (1973). Sample functions of the N-parameter Wiener process. *Ann. Probab.* 1, 138-163.
- [68] S.R. Paranjape, C. Park (1973). Laws of iterated logarithm of multiparameter Wiener processes. *J. Multivariate Anal.* 3, 132-136.
- [69] D. Pollard (1981). Strong consistency of k-means clustering. *Ann. Statist.* 9, 135-140.
- [70] D. Pollard (1982). A central limit theorem for k-means clustering. *Ann. Probab.* 10, 919-926.

- [71] D. Pollard (1984). *Convergence of Stochastic Processes*, Springer-Verlag, New York.
- [72] R. Pyke (1972). Partial sums of matrix arrays and Brownian sheets. E.F. Harding, D.G. Kendall (eds.). *Stochastic Analysis*, Wiley, New York, 331-348.
- [73] R. Pyke, G.R. Shorack (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist.* 39, 755-771.
- [74] F.H. Ruymgaart, J.A. Wellner (1982). *Growth Properties of Multivariate Empirical Processes*, Report 8202, Catholic University, Nijmegen.
- [75] F.H. Ruymgaart, J.A. Wellner (1984). Some properties of weighted multivariate empirical processes. *Statist. Decisions* 2, 199-223.
- [76] G.R. Shorack (1980). Some law of the iterated logarithm type results for the empirical process. *Austral. J. Statist.* 22, 50-59.
- [77] G.R. Shorack, J.A. Wellner (1978). Linear bounds on the empirical distribution function. *Ann. Probab.* 6, 349-353.
- [78] G.R. Shorack, J.A. Wellner (1982). Limit theorems and inequalities for the uniform empirical process indexed by intervals. *Ann. Probab.* 10, 639-652.
- [79] G.R. Shorack, J.A. Wellner (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.
- [80] B.W. Silverman (1983). Convergence of a class of empirical distribution functions of dependent random variables. *Ann. Probab.* 11, 745-751.
- [81] A.V. Skorohod (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* 1, 261-290.
- [82] N.V. Smirnov (1944). Approximate laws of distribution of random variables from empirical data. *Uspehi Mat. Nauk* 10, 179-206 (in Russian).
- [83] W. Stute (1982). The oscillation behavior of empirical processes. *Ann. Probab.* 10, 86-107.
- [84] W. Stute (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab.* 12, 361-379.
- [85] V.N. Vapnik, A.Ya. Červonenkis (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.* 16, 264-280.

- [86] J.A. Wellner (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete* 45, 73-88.
- [87] M.J. Wichura (1969). Inequalities with applications to the weak convergence of random processes with multi-dimensional time parameters. *Ann. Math. Statist.* 40, 681-687.

NOTATION INDEX

A_d , 3	Q_k , 24
β_σ , 48	R , 2
D_d , 3	\bar{R} , 2
$D_{n,v}^{(k_n)}$, 53	$ R $, 18
$D_{n,v}^{(k)}(k_n)$, 53	$R(s,t)$, 2
$ dt $, 15	ρ , 3
F , 2	$<$, 2
\bar{F} , 15	\leq , 2
F_n , 2	$ t $, 6
$[\cdot]$, 10	$\langle t_j \rangle$, 2
I^d , 2	$\langle t_1, \dots, t_d \rangle$, 2
L_d , 3	U , 6
\wedge , 2	U_n , 6
M_n , 67	$U_{n,v}$, 19
ω_n , 67	U_n^\pm , 81
ω'_n , 79	U^F , 3
$\bar{\omega}_n$, 79	U_n^F , 2
ω_n^\pm , 81	$v_{n,v}$, 39
(Ω, F, P) , 2	$v_{n,v}^{(k)}$, 39
ψ , 11	\vec{w} , 3
Q , 13	$x_{k:n}$, 37
Q^* , 5	

SUMMARY

In this thesis multivariate empirical processes based on a sequence of independent and identically distributed random vectors are considered. As indexing sets either quadrants, which are identified with points, or rectangles with sides parallel to the coordinate axes are used. In this setup a number of optimal results is obtained concerning weak convergence and strong limit theorems; in most of these results the empirical processes are weighted.

In Chapter 1 a brief introduction, including the main notation, is given and some basic facts are presented. Chapter 2 is devoted to probability inequalities. In section 1 a probability inequality for the local behaviour of the empirical process is derived with the aid of Poissonization; there the distribution function is completely arbitrary. In the second section this inequality becomes a fundamental tool for deriving global probability inequalities for weighted empirical processes under the assumption that the underlying distribution is uniform on the unit "square". In that section also some maximal inequalities are derived.

In Chapter 3 necessary and sufficient conditions on the weight functions are established for weak convergence of the weighted multivariate uniform empirical processes, indexed by points and rectangles respectively. The proper weighting is described for small as well as large sets in the indexing classes. Chapter 4 is devoted to strong limit theorems for suprema of the absolute values of the weighted multivariate uniform empirical processes for special choices of the weight functions. There the processes are mostly indexed by points since indexation by rectangles is not very meaningful. The aforementioned suprema are taken over various types of regions: the whole unit "square", its middle and its tail. It turns out that in some cases laws of the iterated logarithm can be established like in the unweighted case, whereas in other situations no proper standardization is possible. Chapter 5 is concerned with results for two types of oscillation moduli of the unweighted uniform empirical process. More specifically in section 1 the order of magnitude of the moments of these

oscillation moduli is obtained for "sufficiently large" sample size and in section 2 a complete picture of the almost sure behaviour of one of these moduli is given.

Finally in Chapter 6 some generalizations and potential applications are considered. In section 1 a number of results for the negative parts of the uniform empirical process is presented. In section 2 the assumption is dropped that the underlying distribution is uniform. There the case that the empirical processes are based on random vectors from a distribution having a continuous density that is bounded away from 0 and ∞ is briefly considered. It should be emphasized that again the local probability inequality (Chapter 2, section 1) is the fundamental tool for the aforementioned results. A few miscellaneous results, i.a. for "other" empirical processes, are discussed in section 3. In conclusion, the following fields of possible applications are mentioned in section 4: extreme value theory, U-statistics, general spacings and density and regression function estimation.

SAMENVATTING

MULTIVARIATE EMPIRISCHE PROCESSEN

In dit proefschrift worden multivariate empirische processen, gebaseerd op een rij van onafhankelijke en identiek verdeelde stochastische vectoren, beschouwd. Als indexverzamelingen worden kwadranten, die geïdentificeerd worden met punten, of rechthoeken met zijden parallel aan de coördinaatassen gebruikt. In deze opzet wordt een aantal optimale resultaten betreffende zwakke convergentie en sterke limietstellingen verkregen; in de meeste van deze resultaten zijn de empirische processen gewogen.

Behalve een korte introductie en de voornaamste notatie worden in Hoofdstuk 1 enige bekende resultaten gepresenteerd. Hoofdstuk 2 is gewijd aan kansongelijkheden. In paragraaf 1 wordt een kansongelijkheid voor het lokale gedrag van het empirische proces afgeleid met behulp van Poissonisering; in die paragraaf is de verdelingsfunctie volkomen willekeurig. In de tweede paragraaf worden m.b.v. deze ongelijkheid globale kansongelijkheden afgeleid voor gewogen empirische processen onder de voorwaarde dat de onderliggende verdeling uniform is op het "eenheidsvierkant". In die paragraaf worden ook enige maximale ongelijkheden afgeleid.

In Hoofdstuk 3 worden nodige en voldoende voorwaarden op de gewichtsfuncties bepaald voor zwakke convergentie van de gewogen multivariate uniforme empirische processen, geïndiceerd door respectievelijk punten en rechthoeken. De geschikte weging is beschreven voor kleine alsmede grote verzamelingen uit de indexklassen. Hoofdstuk 4 is gewijd aan sterke limietstellingen voor suprema van de absolute waarden van de gewogen multivariate uniforme empirische processen voor speciale keuzen van de gewichtsfuncties. In dat hoofdstuk worden de processen meestal geïndiceerd door punten aangezien indicering door rechthoeken niet erg zinvol is. De voornoemde suprema worden genomen over verscheidene typen gebieden: het hele "eenheidsvierkant", het midden en de staart. In sommige gevallen blijkt dat wetten

van de geïtereerde logaritmie bewezen kunnen worden, net als in het ongewogen geval. In andere situaties echter, blijkt geschikte standaardisatie onmogelijk te zijn. In Hoofdstuk 5 worden resultaten voor twee typen van oscillatiemoduli van het ongewogen uniforme empirische proces afgeleid. In paragraaf 1 wordt de orde van grootte van de momenten van deze oscillatiemoduli verkregen voor "voldoend grote" steekproefomvang en in paragraaf 2 wordt een compleet overzicht van het "almost sure" gedrag van een van deze moduli gegeven.

In Hoofdstuk 6, tenslotte, worden enige generalisaties en mogelijke toepassingen beschouwd. In paragraaf 1 wordt een aantal resultaten voor de negatieve delen van het uniforme empirische proces gepresenteerd. In paragraaf 2 wordt de voorwaarde dat de onderliggende verdeling uniform is verlaten. Daar worden kort empirische processen bekeken die gebaseerd zijn op stochastische vectoren uit een verdeling met een continue en van 0 en ∞ weggegrensde dichtheid. Het is interessant om op te merken dat voornoemde resultaten weer m.b.v. de lokale kansongelijkheid (Hoofdstuk 2, paragraaf 1) afgeleid worden. Een paar gemengde resultaten, o.a. voor "andere" empirische processen, worden besproken in paragraaf 3. Tot slot worden de volgende gebieden van mogelijke toepassingen genoemd in paragraaf 4: extreme waarde theorie, "U-statistics", algemene "spacings" en het schatten van dichtheden en regressiefuncties.

CURRICULUM VITAE

John H.J. Einmahl werd op 14 februari 1957 te Heerlen geboren. Hij behaalde in 1975 aan het Sophianum te Vaals zijn diploma Atheneum-B. Hij studeerde vervolgens wiskunde aan de Katholieke Universiteit Nijmegen en legde in september 1982 zijn doctoraalexamen af met als hoofdrichting grondslagen van de wiskunde. In november van dat jaar legde hij een aanvullend doctoraalexamen in mathematische statistiek af. Gedurende het jaar 1981 was hij tevens wiskundeleraar aan het Elzendaalcollege te Boxmeer. Van oktober 1982 tot april 1986 was hij werkzaam als wetenschappelijk medewerker aan de Katholieke Universiteit Nijmegen. Aldaar verrichtte hij onderzoek op het gebied van de multivariate empirische processen onder leiding van prof.dr. F.H. Ruymgaart, hetgeen leidde tot dit proefschrift. Vanaf april 1986 is hij werkzaam als universitair docent aan de Rijksuniversiteit Limburg.